

**This Page Is Inserted by IFW Operations
and is not a part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- **BLACK BORDERS**
- **TEXT CUT OFF AT TOP, BOTTOM OR SIDES**
- **FADED TEXT**
- **ILLEGIBLE TEXT**
- **SKEWED/SLANTED IMAGES**
- **COLORED PHOTOS**
- **BLACK OR VERY BLACK AND WHITE DARK PHOTOS**
- **GRAY SCALE DOCUMENTS**

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 8, 2003, 00:13:01 ; Search time 51 Seconds
(without alignments)
5390.479 Million cell updates/sec

Title: US-10-037-270-482

Perfect score: 1560

Sequence: 1 ggcacttaccctcctgagcta.....tgggaatcagacatactagt 866

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 2215726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ n2p.model -DEV=xlh
-Q=/cgn2_1/USFTO_epool/US10037270/runat_07112003_140513_8736/app_query_faeta_1.1031
-DB=A.Geneseq_15jun03 -OFMT=faeta -SUFFIX=rag -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cd1
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10037270 @cgn 1.1 64 @runat 07112003_140513_8736 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSFBLOCK=100 -LONGCLOG
-DEV_TIMEOUT=120 -MAPN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1298	83.2	248	22	AAW41224
2	1255	80.4	235	22	AAW39438
3	1255	80.4	235	23	AAW47664
4	1255	80.4	235	23	AAW47665
5	1250	80.1	235	23	AAW76316
6	1247	79.9	235	22	AAU07697
7	1247	79.9	235	22	AAU07696
8	1247	79.9	235	23	AAU74746
9	1247	79.9	235	24	AAU873741
10	1222	78.3	265	23	AAU82744
11	1216	77.9	228	23	AAW47674
12	1216	77.9	228	23	AAW47674
13	1216	77.9	228	23	AAW47674
14	1216	77.9	228	23	AAW47674
15	1216	77.9	228	23	AAW47674
16	1216	77.9	228	23	AAW47674
17	1216	77.9	228	23	AAW47674
18	1216	77.9	228	23	AAW47674
19	1216	77.9	228	23	AAW47674
20	1216	77.9	228	23	AAW47674
21	1216	77.9	228	23	AAW47674
22	1216	77.9	228	23	AAW47674
23	1216	77.9	228	23	AAW47674
24	1216	77.9	228	23	AAW47674
25	1216	77.9	228	23	AAW47674
26	1216	77.9	228	23	AAW47674
27	1216	77.9	228	23	AAW47674
28	1216	77.9	228	23	AAW47674
29	1216	77.9	228	23	AAW47674
30	1216	77.9	228	23	AAW47674
31	1216	77.9	228	23	AAW47674
32	1216	77.9	228	23	AAW47674
33	1216	77.9	228	23	AAW47674
34	1216	77.9	228	23	AAW47674
35	1216	77.9	228	23	AAW47674
36	1216	77.9	228	23	AAW47674
37	1216	77.9	228	23	AAW47674
38	1216	77.9	228	23	AAW47674
39	1216	77.9	228	23	AAW47674
40	1216	77.9	228	23	AAW47674
41	1216	77.9	228	23	AAW47674
42	1216	77.9	228	23	AAW47674
43	1216	77.9	228	23	AAW47674
44	1216	77.9	228	23	AAW47674
45	1216	77.9	228	23	AAW47674

ALIGNMENTS

RESULT 1
AAW41224
AAW41224 standard; Protein; 248 AA.
XX
XX
AC AAW41224;
XX
XX
DT 22-OCT-2001 (first entry)
XX
XX
DE Human polypeptide SEQ ID NO 6155.
XX
XX
KW Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
peripheral nervous system; neuropathy; central nervous system; CNS;
Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;
amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemocarcin;
chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
leukaemia.
XX
XX
OS Homo sapiens.

XX MO200153312-A1.
XX 26-JUL-2001.
XX 26-DEC-2000; 2000MO-US34263.
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0662191.
XX 19-OCT-2000; 2000US-0693036.
XX 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.
XX Tang Y, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI, 2001-442253/47.
XX N-PSDB; AAI60380.
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX Example 2; SEQ ID NO 6155; 10078bp; English.
XX The invention relates to human nucleic acids (AA157798-AA161369) and
XX the encoded polypeptides (AAM38642-AAM42213) with nootropic.
XX immunosuppressant and cyostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders.
XX Note: The sequence data for this patent did not form part of the printed
XX specification.
XX SQ Sequence 248 AA;
XX Alignment Scores:
XX Pred. No.: 3,53e-131 Length: 248
XX Score: 1298.00 Matches: 244
XX Percent Similarity: 98.39% Conservative: 0
XX Best Local Similarity: 98.39% Mismatches: 4
XX Query Match: 83.21% Indels: 0
XX DB: 22 Gaps: 0
XX US-10-037-270-482 (1-866) x AAM41224 (1-248)
XX 3 CACTTACTCCCTGAGCTAAGGGGAGAGCTGATCACCAGTAATATGCTTCTATTG 62
XX 1 HisleuLeuProGluLeuAArgGlyLysSerTrrplehnehtcystryValPheIyLeu 20
XX 63 GGATGCTCGCTGGGACATTTTCTTGTGCTGACTCATCTGTTGAGAAAGAACCTGCT 122
XX 21 GlyValLeuAaGlyThrPhePhePheAlaAspSerSerValGlnLysGluAspProAla 40
XX 123 CCGATTTGTTGTAACCCAGTGTCACTCAACCCCGTGTGGCGGCTCCATCAAAACC 162
XX 41 ProIyLeuValIyLeuLysSerHisPheAsnProCyValGlyValLeuIleLysPro 60
XX 183 AGCTGGGTGCTGGCCCGAGCTCACTGATTTACCAATGTGAAGTATGCTGGAAAT 242

DB 61 SerTrpValLeuAlaProAlaHisCystryLeuProAsnLeuLysValMetLeuGlyAsn 80
QY 243 TTCAGAGAGAGAGTACAGAGCGGTACTGAAAGACAAATTAACCCATTGATGTCGCGC 302
DB 81 PheLysSerArgValAlaArgAspGlyThrGluGlnThrIleAsnProIleGlnIleValArg 100
QY 303 TACTGAACTACAGTCATAGCCGCCCAAGATGATGATGCTCATCAAGTGGCTAAG 362
DB 101 TyTrpAsnIySerHisSerIleAlaProGlnAspPheMetLeuIleLysValLys 120
QY 363 CCTGGCATGTCTCATCTCCAAAGTCCAGCCCTTCCCTGCCACCAATGTCAGGCCA 422
DB 121 ProAlaMetLeuAsnProLysValGlnAlaLeuAsnProProThrThrAsnValArgPro 140
QY 423 GGCACGTCTGTCTACTCTCAGGTTTGATGTCGAGCCAGAAAGAGTGGCGACACCT 482
DB 141 GlyThrValCysLeuLeuSerGlyLeuAspTrpSerGlnGluAsnSerGlyArgHisPro 160
QY 483 GACTTCGCGCAGAACTGAGAGCGCCCGCTGATGTCTGATGAGAAATGCCAAAAACAGAA 542
DB 161 AspLeuAArgGlnAsnLeuGlnAlaProValMetSerAspArgGlyGlyGlnIystrGlu 180
QY 543 CAAAGAAAAAGCCACAGAAATTCCTTATGTGTGAATTTGTGAAGATTACGCCAATT 602
DB 181 GlnGlyLysSerHisArgAsnSerLeuCyValIyPheValIyPheSerArgIle 200
QY 603 TTTGGGAGGTGGCGCTGCTGCTGATCTGCAAGAACAGCTCCAGGAAATCGAGGTG 662
DB 201 PheGlyGluValAlaValAlaThrValIleCyLysAspLysLeuGlnIyIleGluVal 220
QY 663 GGGCACTTCATGAGGAGGAGCGTGGCATCTACCAATGTTTCAAAATATGATCTGCG 722
DB 221 GlyHisPheMetGlyGlyAspValGlyIleTyThrAsnValIyTrpIyValSerTrp 240
QY 723 ATTGAGAACACTGCTTAAGACAG 746
DB 241 IleGluAsnThrAlaLysAspLys 248
RESULT 2
AAM39438
ID AAM39438 standard; Protein; 235 AA.
XX AAM39438;
XX 22-OCT-2001 (first entry)
XX Human polypeptide SEQ ID NO 2583.
XX Human; nootropic; immunosuppressant; cyostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia.
XX Homo sapiens.
XX OS
XX PN WO200153312-A1.
XX 26-JUL-2001.
XX 26-DEC-2000; 2000MO-US34263.
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0662191.
XX 19-OCT-2000; 2000US-0693036.
XX 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI; 2001-442253/47.
 DR N-PSDB; AA158594.
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Example 4; SEQ ID NO 2583; 10078bp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AA158642-AA162213) with neurotropic,
 CC immunosuppressant and cyrostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Actvlin/inhibin activity, chemoclastic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 CC
 XX
 XX
 SQ Sequence 235 AA;
 Alignment Scores:
 Pred. No.: 1.53e-126 Length: 235
 Score: 1255.00 Matches: 235
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 80.45% Indels: 0
 DB: 22 Gaps: 0
 US-10-037-270-482 (1-866) x AAM39438 (1-235)
 QY 42 ATGAATATGCTCTTATTTGGTGTCTGCTGGGACATTTTCTTGTGACTCATCT 101
 DB 1 MetLysTyrValIhePheTyrLeuGlyValLeuAlaGlyThrPhePheAlaAspSer 20
 QY 102 GTTCAGAAAGAGACCCCTGCTCCCTATTGGTGTACCTCAAGTCTCACTTCAACCCCTGT 161
 DB 21 ValGlnLysGlnAspProAlaProTyrLeuValTyrLeuLysSerHisPheAsnProCys 40
 QY 162 GTGGGGCTCTCATCAAAACCACTGGGTGCTGCCCCAGCTCACTGCTTTTACCAAT 221
 DB 41 ValGlyValLeuLysProSerTrpValLeuAlaProAlaHisCysTyrLeuProAsn 60
 QY 222 CTGAAGTGTGCTGGGAATTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 281
 DB 61 LeuLysValMetLeuLysAsnLysPheSerArgValAlaGlnAspLysGlnThrIle 80
 QY 282 AACCCATTGATGCTGCTCGCTACTAGTAATCAAGTATGACCCGACAGATGACCTC 341
 DB 81 AsnProLysGlnLysValArgTyrTrpAsnTyrSerHisSerAlaProGlnAspLys 100
 QY 342 ATGCTCATCAAGTGTGCTAAGCTGCGCATCTCAATCCCAAGTCCAGCCCTTCCCTC 401
 DB 101 MetLeuLysValLeuLysAlaLysProAlaMetLeuAsnProLysValGlnProLeuProLeu 120
 QY 402 GCCACCAACATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 461
 DB 121 AlaThrThrAsnValAlaGlyProGlyThrValCysLeuLysSerLysLeuAspTrpSerGln 140
 QY 462 GAAAGAGTGGCGGACCACTGACTTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 521
 DB 141 GluAsnSerLysArgHisProAspLeuArgGlnAsnLeuLysAlaProValMetSerAsp 160

QY 522 CGAATGCCAAAAACAGAACAGAAAGAACAGAAATTCCTTATGTGAATTT 581
 DB 161 ArgGluCysGlnLysPheThrLysGlnLysSerHisArgAsnSerLeuValLysPhe 180
 QY 582 GTGAAGATATTCAGCCGAATTTTGGGAGGTGGCCGTGCTACTGCTATCGCAAGAC 641
 DB 181 ValLysValIlePheSerArgIlePheGlyValAlaValAlaThrValIleCysLysAsp 200
 QY 642 AAGCTCCAGGAATGAGGTGGGACCTTCATGCGAGGGAGAGCTCGGCATCTACCAAT 701
 DB 201 LysLeuGlnGlyIleGluValAlaGlyHisPheMetGlyLysPvalGlyIleTyrThrAsn 220
 QY 702 GTTACAAATATGATCTCTGATGAGATGAGACACTGCTAAGACACAG 746
 DB 221 ValTyrLysTyrValSerTrpIleGluAsnThrAlaLysAspLys 235
 RESULT 3
 AAM47664
 ID AAM47664 standard; Protein; 235 AA.
 XX
 AC AAM47664;
 XX
 DT 21-FEB-2002 (first entry)
 XX
 DE MOL6a protein sequence.
 XX
 KW MOL; G-coupled protein-receptor; cardiomyopathy; atherosclerosis;
 KW cell signal processing; metabolic disorder; diabetes; cancer;
 KW neurodegenerative disorder; immune disorder; cardiac disorder;
 KW lung disease; autoimmune disease; developmental disorder; antidiabetic;
 KW Cytostatic; Neuroprotective; Antiatherosclerotic; Immunosuppressive;
 KW Gene therapy; Vaccine; antiinflammatory; MOL6a; trypsin-like.
 XX
 OS Unidentified.
 XX
 PN MO200181578-A2.
 XX
 PD 01-NOV-2001.
 XX
 XX
 XX 26-APR-2001; 2001WO-US13578.
 XX
 XX 26-APR-2000; 2000US-200158P.
 XX
 PR 26-APR-2000; 2000US-200613P.
 PR 28-APR-2000; 2000US-200780P.
 PR 28-APR-2000; 2000US-200780P.
 PR 01-MAY-2000; 2000US-201006P.
 PR 01-MAY-2000; 2000US-201007P.
 PR 01-MAY-2000; 2000US-201236P.
 PR 01-MAY-2000; 2000US-201238P.
 PR 02-MAY-2000; 2000US-201186P.
 PR 03-MAY-2000; 2000US-201474P.
 PR 03-MAY-2000; 2000US-201508P.
 PR 25-JUL-2000; 2000US-220591P.
 PR 15-SEP-2000; 2000US-232678P.
 PR 22-JAN-2001; 2001US-263217P.
 PR 30-JAN-2001; 2001US-265160P.
 XX
 XX (CURA-) CURAGEN CORP.
 PA
 XX Vermet CAM, Fernandes ER, Gerlach V, Shinkets RA, Malysankar UM;
 PI Boldet FL, Zehnen BD, Spytek KA, Majumder K, Tchernav VT;
 PI Radisaru M, Patturajan M, Burgess CE, Gangoli BA, Smithson G;
 PI Rastelli L, Macdougall JR, Taupier KO, Grose WM, Szekeres BS;
 PI Alsbrook JP;
 XX
 DR WPI; 2002-049278/06.
 DR N-PSDB; ABA04594.
 XX
 PT Novel G-protein coupled receptor-related polypeptides and
 PT polynucleotides for diagnosing, preventing and treating cardiomyopathy,
 PT atherosclerosis, disorders related to cell signal processing and for
 PT identifying modulators -

PS Claim 1; Page 44; 227bp; English.

XX The present invention relates to novel G-coupled protein-receptor related
 CC proteins and coding sequences (MOLX, where X is a number from 1 to 10,
 CC ABA04589-ABA04603 and AAM47659-AAM47673). MOLX proteins and coding
 CC sequences are useful for treating or preventing a MOLX-associated
 CC disorder, such as cardiomyopathy, atherosclerosis, disorders related to
 CC cell signal processing and metabolic pathway modulation, diabetes and
 CC cancer. Additionally, MOLX proteins and coding sequences are useful for
 CC preventing and treating a variety of disorders including metabolic
 CC disorders, nutritional oedema, chronic and hereditary pancreatitis,
 CC obesity, infectious disease, anorexia, neurodegenerative disorders,
 CC Alzheimer's disease, Parkinson's disease, stroke, immune disorders,
 CC haematopoietic disorders and various dyslipidemias, metabolic syndrome X
 CC and wasting disorders associated with chronic diseases and cancers,
 CC cardiac disorders, hypertension, hypercalcaemia, cirrhosis, angiogenesis
 CC and wound healing, trauma, glomerulonephritis, hyper and hypothyroidism,
 CC multiple sclerosis, lung diseases including asthma, Crohn's disease,
 CC scleroderma, autoimmune diseases, developmental disorders and neural tube
 CC defects. The present sequence is the protein sequence for MOL6a.
 CC MOL6a is a trypsin-like protein.

XX Sequence 235 AA;

Alignment Scores:
 Pred. No.: 1.53e-126 Length: 235
 Score: 1255.00 Matches: 235
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 80.45% Indels: 0
 DB: 23 Gaps: 0

US-10-037-270-482 (1-866) x AAM47664 (1-235)

QY 42 ATGAAATATGCTCTTCTATTGGGCTGCTGCTGGGACATTTTCTTCTGCTACTCT 101
 DB 1 MetLysrYrValPheYrLLeuGlyValLeuAlaGlyThrPhePheAlaAspSer 20
 QY 102 GTTCAGAAAGAGACCTGCTCCCTATTGGTGTAGCTCAAGTCTCACTTCAACCCCTGT 161
 DB 21 ValGlnLysGlnuSprProLapProYrLLeuValYrLeuLysSerHisPheAsnProCys 40
 QY 162 GTGGGGCTCTCTCAACACCCAGCTGGGTCTGGCCCCAGCTCACTGCTATTACCAAT 221
 DB 41 ValGlyValLeuLLeuLysProSerTrpValLeuAlaProLahisCysYrLeuProAsn 60
 QY 222 CTGAAAGTGTGCTGGGAAATTTCAAGACAGAGTCAAGACGGTATGACACACAT 281
 DB 61 LeuLysValMetLeuGlnLysAsnPheLysSerArgValArgAspGlyThrGlnThrIle 80
 QY 282 AACCCCATTCAGATCGTCCGCTACTGGAATCACTAGTATGCGCCCAAGATGACCTC 341
 DB 81 AsnProLLeuGlnLLeuValArgYrTrpAsnTrpSerHisSerAlaProGlnAspLeu 100
 QY 342 ATGCTCATCAGCTGGCTAAAGCTGCATGCTCAATCCCAAGTCCAGCCCTTCCCTC 401
 DB 101 MetLeuLLeuLysLeuAlaLysProLahMetLeuAsnProLysValGlnProLeuProLeu 120
 QY 402 GCCACCACTCATGTCAAGGCAAGGCAAGCTGTCTGTATCTCAAGTTGAGCTGAGCCAA 461
 DB 121 AlaThrThrAsnValArgProGlyThrValCysLeuLeuSerGlyLeuAspTrpSerGln 140
 QY 462 GAAACAGTGGCGGACCCCTGACTGGCAGAACTGAGAGGCCCGCGAGTGTGAT 521
 DB 141 GlnAsnSerGlyArgHisProAspLeuArgGlnAsnLeuGlnuAlaProValMetSerAsp 160
 QY 522 CGAAGATGCCAAAAACAAGAAAGAAAGCAAGCAAGATTCCTTATGTGAAATTT 581
 DB 161 ArgGlnCysGlnLysThrGlnGlnGlyLysSerHisArgHisSerLeuCysValLysPhe 180
 QY 582 GTGAAAGTATTCAGCGCAATTTTGGGAGGTGGCCGCTCTCATCTGCAACAAAC 641
 DB 181 ValLysValPheSerArgLLeuPheGlyGlnuValAlaValAlaThrValLLeuCysLysAsp 200

QY 642 AACCTCAGAGGAATTCAGATGGGCGCACTTCATGAGAGGAGGACCTGCGCATCTACCAAT 701
 DB 201 LysLeuGlnLLeuLLeuValGlnHisPheMetClyGlyAspValGlyLLeuYrThrAsn 220
 QY 702 GTTTACAAATATGATTCCTGGATTGAGAACACTGCTTAGAGCAAG 746
 DB 221 ValTrpLysrYrValSerTrpLLeuGlnuAsnThrAlaLysAspLys 235

RESULT 4

AAM47665
 ID AAM47665 standard; Protein; 235 AA.

XX AAM47665;

DT 21-FEB-2002 (first entry)

DE MOL6b protein sequence.

XX MOL; G-coupled protein-receptor; cardiomyopathy; atherosclerosis;
 KM cell signal processing; metabolic disorder; diabetes; cancer;
 KM neurodegenerative disorder; immune disorder; cardiac disorder;
 KM lung disease; autoimmune disease; developmental disorder; antidiabetic;
 KM Cytostatic; Neutroprotective; Antiatherosclerotic; Immunosuppressive;
 KM Gene therapy; Vaccine; antiinflammatory; MOL6b; trypsin-like.

XX Unidentified.

OS W0200181578-A2.

PN 01-NOV-2001.

XX 26-APR-2001; 2001MO-US13578.

XX 26-APR-2000; 2000US-200158P.

PR 28-APR-2000; 2000US-200613P.

PR 28-APR-2000; 2000US-200780P.

PR 01-MAY-2000; 2000US-201006P.

PR 01-MAY-2000; 2000US-201007P.

PR 01-MAY-2000; 2000US-201236P.

PR 01-MAY-2000; 2000US-201238P.

PR 02-MAY-2000; 2000US-201186P.

PR 03-MAY-2000; 2000US-201474P.

PR 03-MAY-2000; 2000US-201508P.

PR 25-JUL-2000; 2000US-220591P.

PR 15-SEP-2000; 2000US-232678P.

PR 22-JAN-2001; 2001US-263217P.

PR 30-JAN-2001; 2001US-265160P.

XX (CURA-) CURAGEN CORP.

XX Vermont CAM, Fernandes ER, Gerlach V, Shinkete RA, Majumkar UM;

PI Bolding FL, Zernhuken BD, Spytek KA, Majumder K, Tcherney VT;

PI Padigaru M, Patlurajan M, Burgess CE, Gangolli BA, Smithson G;

PI Rastelli L, Macdougall JR, Taupier RJ, Grose WM, Szekeres ES;

PI Alsobrook JP;

DR WPI: 2002-049278/06.

DR N-PSDB; ABA04595.

XX Novel G-protein coupled receptor-related polypeptides and

PT polynucleotides for diagnosing, preventing and treating cardiomyopathy,

PT atherosclerosis, disorders related to cell signal processing and for

PT identifying modulators -

PS Claim 1; Page 45; 227bp; English.

XX The present invention relates to novel G-coupled protein-receptor related
 CC proteins and coding sequences (MOLX, where X is a number from 1 to 10,
 CC ABA04589-ABA04603 and AAM47659-AAM47673). MOLX proteins and coding
 CC sequences are useful for treating or preventing a MOLX-associated
 CC disorder, such as cardiomyopathy, atherosclerosis, disorders related to
 CC cell signal processing and metabolic pathway modulation, diabetes and

CC infarction), blood vessel (e.g. atherosclerosis, hypertension, Kawasaki
 CC syndrome), and platelet disorders, as well as disorders associated with
 CC bone metabolism (e.g. osteoporosis, osteomalacia, osteopenia, tropical
 CC sprue, rickets), haematopoietic (e.g. diabetes mellitus, rheumatoid
 CC arthritis, autoimmune thyroiditis), liver disorders, viral diseases and
 CC pain or metabolic disorders (e.g. obesity, anorexia, cachexia) and
 CC many other diseases and disorders given in the specification. The
 CC present sequence is the serine protease 14087.

XX Sequence 235 AA;

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:
Score: 1250.00	235	234	1
Percent Similarity: 100.00%		Mismatches: 0	
Best Local Similarity: 99.57%		Indels: 0	
Query Match: 80.13%		Gaps: 0	
DB: 23			

US-10-037-270-482 (1-866) x AAU07697 (1-235)

```

QY 42 ATGAATATGCTCTTATTGGGTCCTGCTGGGACATTTTCTTGGTCACTCATCT 101
DB 1 MetLysTyrValPheTyrLeuGlyValLeuAlaGlyThrPhePheAlaAspSerSer 20
QY 102 GTTCAGAAAGAGACCTGCTCCCTATTGGTGTACCTCAAGTCTCACTTCAACCCCTGT 161
DB 21 ValGlnLysGluAspProAlaProTyrLeuValTyrLeuLysSerHisPheAsnProCys 40
QY 162 GTGGGCGCTCTCATCAACCCAGCTGGGTCGTGGCCCGGCTCACTGCTATTATTCACAAT 221
DB 41 ValGlyValLeuLeuLysProSerTyrPValLeuAlaProAlaHisCysTyrLeuProAsn 60
QY 222 CTGAAGTGAATGCTGGGAATTTCAAGAGAGAGTCAAGAGCGGTACTGGAACAGCAAT 281
DB 61 LeuLysValMetLeuGlyAsnPhelYserArgValArgAspGlyThrGlnThrIle 80
QY 282 AACCCCATTCAGATCGTCCGCTACTGGAAGTCAAGTACGCGCCAGAGATGATCCTC 341
DB 81 AsnProLeuGlnLeuValArgTyrTyrPheSerHisPheAlaProGlnAspLeu 100
QY 342 ATGCTCATCAAGCTGGCTAGACCTGCAATGCTCAATCCCAAGTCCAGCCCTTCCCTC 401
DB 101 MetLeuLysLeuAlaLysProAlaMetLeuHisPheLysValGlnProLeuProLeu 120
QY 402 GCCACCAACATGTCAGGCCAGGACGCTGCTTACTTCTCAAGTTTGAAGTGAAGCAA 461
DB 121 AlaThrThrAsnValArgProGlyThrValCysLeuLeuSerGlyLeuAspTyrSerGln 140
QY 462 GAAACAGTGGCCGACACCTGACTGCGGAGAACTGAGAGGCCCGGTGATGATGAT 521
DB 141 GluAsnSerGlyArgHisPheAspLeuArgGlnAsnLeuGlnAlaProValMetSerAsp 160
QY 522 CGAGAAATGCCAAAAACAGAAACAAGAAAAAGCCAGAGAAATTCCTATGTGAAATTT 581
DB 161 ArgGlnCysGlnLysThrGlnGlnGlyLysSerHisArgAsnSerLeuCysValLysPhe 180
QY 582 GTGAAGTATTCACCCGAATTTTGGGAGGTGCGGTGCTATGTCATCTGCAAGAC 641
DB 181 ValLysValPheSerArgIlePheGlyGluValAlaValAlaThrValIleCysValAsp 200
QY 642 AAGCTCAGAGGAATCGAGTGGGACACTTCATGAGAGGAGGAGTCCGATCTACCAAT 701
DB 201 LysLeuGlnGlyIleGluValGlyHisPheMetCysLysPheValGlyIleTyrThrAsn 220
QY 702 GTTACAAATATGATCTCTGATTGAGAAACACTGTAAGCAAG 746
DB 221 ValTyrLysTyrValSerTyrIleGluAsnThrAlaLysAspLys 235

```

RESULT 6

AAU07697
 ID AAU07697 standard; Protein; 235 AA.

```

AC AAU07697;
XX 18-DEC-2001 (first entry)
DT Human Ztcrp3 serine protease polypeptide.
XX
DE Human Ztcrp3 serine protease polypeptide.
XX
XX Human; Ztcrp3; serine protease; asthma; vascular function; inflammation;
XX gene therapy; stroke; testicular function; spermatogenesis; haemostatic;
XX mass spectrometry; circular dichroism; X-ray crystallography;
XX nuclear magnetic resonance spectroscopy; antiasthmatic; antiinflammatory;
XX cerebroprotective.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..19
XX Protein /note="Signal peptide"
XX /note="Mature human Ztcrp3"
XX
XX WO200166771-A2.
XX
XX 13-SEP-2001.
XX
XX 28-FEB-2001; 2001MO-US06432.
XX
XX 03-MAR-2000; 2000US-0518387.
XX
XX (ZYMO ) ZYMOGENETICS INC.
XX
XX Conk11n DC;
XX
XX WPI; 2001-589946/66.
XX
XX DR N-PSDB; AAS12970.
XX
XX PT Novel Ztcrp3 polypeptides and polynucleotides useful in the treatment
XX of asthma, vascular disorders including stroke, inflammation and
XX testicular function
XX
XX PS Claim 3; Page 2-3; 82pp; English.
XX
XX The invention relates to an isolated human Ztcrp3 polypeptide, a member
XX of the serine protease family. Ztcrp3 polypeptides and their associated
XX polynucleotides are useful in diagnosis, therapy and industry and are
XX used as targets for identifying modulators, preferably inhibitors of
XX serine protease activity. The sequences are useful in the treatment of
XX asthma, vascular function such as stroke, inflammation and testicular
XX function (by modulating spermatogenesis). Ztcrp3 proteins can be used for
XX identifying peptide cleavage sites and for coupling amino and carboxy
XX terminal tags. The polypeptides are also useful to teach analytical
XX skills such as mass spectrometry, circular dichroism, X-ray
XX crystallography and nuclear magnetic resonance spectroscopy. This
XX sequence represents a human Ztcrp3 polypeptide.
XX
XX Sequence 235 AA;
XX
XX Alignment Scores:
XX Pred. No.: 1,126-125 Length: 235
XX Score: 1247.00 Matches: 234
XX Percent Similarity: 99.57% Conservative: 0
XX Best Local Similarity: 99.57% Mismatches: 1
XX Query Match: 79.94% Indels: 0
XX DB: 22 Gaps: 0
XX
XX US-10-037-270-482 (1-866) x AAU07697 (1-235)
XX
QY 42 ATGAATATGCTCTTATTGGGTCCTGCTGGGACATTTTCTTGGTCACTCATCT 101
DB 1 MetLysTyrValPheTyrLeuGlyValLeuAlaGlyThrPhePheAlaAspSerSer 20
QY 102 GTTCAGAAAGAGACCTGCTCCCTATTGGTGTACCTCAAGTCTCACTTCAACCCCTGT 161
DB 21 ValGlnLysGluAspProAlaProTyrLeuValTyrLeuLysSerHisPheAsnProCys 40

```

QY	162	GTGGGGCTCCTCATCAAAACCCAGCTGGGGTGGCCCAAGCTCATCTGGTATTTCACAAAT	221
Db	41	ValGlyValIleuIleIysProSerTrpValIleuAlaProIleAsnIleCysTrpIleuProAsn	60
QY	222	CTGAAGATGATGCTGGGAAATTTCAAGACAGAGCTCAGAGACGGTACTGGAACAGACAAATT	281
Db	61	IleuIysValIleMetIleuIleuIleuIysnPhelysSerArgValArgAspGlyThrGluGlnThrIle	80
QY	282	MAACCCATTCAAGTGTCCGCTACTGGAACCTACGTCATACGCGCCCAACAGATGACCTC	341
Db	81	AsnProIleGlnIleValArgTrpIleTrpAsnTrpSerHisSerAlaProGlnAspAspIleu	100
QY	342	ATGCTCATCAAGCTGGCTAAGCGCTGCAGCTGCATATCCCAAGTCACAGCCCTCCCTC	401
Db	101	MetIleuIleIysIleuIleuAlaIysProAlaMetIleuAsnProIysValGlnProIleuThrIleu	120
QY	402	GCCACACCAATATGCTCAGGCCAGGACCATGTGTCTTACTCTCAGATTGGACTGGAGCCAA	461
Db	121	AlaThrThrAsnValArgProGlyThrValCysIleuIleuSerGlyIleuAspTrpSerGln	140
QY	462	GAAGAAAGTGGCCGAGACCCCTGACTTGGGCGAGAACTTGGAGGCCCGGTGATGTCTGAT	521
Db	141	GlnAsnSerGlyArgHisIleProAspIleuArgGlnAsnIleuGlnAlaProValMetSerAsp	160
QY	522	CGAAGATGCAAAAAACAGACACAGAAAAACCCAGAGAAATTCCTTATGTGCAATTT	581
Db	161	ArgGluCysGlnIleIysThrGluGlnIleuIysSerHisAsnGlnSerIleCysValIysPhe	180
QY	582	GTGAAGATTCAGCCGAAATTTTGGGGAGTGGCCGCTGTGCTACTGTTCATCTGCAGAGAC	641
Db	181	ValIysValIlePheSerArgIlePheGlyIleValAlaIleValAlaIleThrValIleCysIysAsp	200
QY	642	AACTCTCCAGGAATCGAGGTGGGCGACTTCATCGGAGGGGACGTTCGCATTTACCAAT	701
Db	201	IysIleuGlnGlyIleGluValGlnIleHisPheMetGlyGlyAspValGlyIleTrpThrAsn	220
QY	702	GTTTACAATATGTATCTCGATTGGAACACTGCTTAGAGACAAG	746
Db	221	ValTrpIleIysTrpValSerTrpIleGluIleuThrAlaIysAspIys	235
RESULT 7			
AAU27666	AAU27666 standard; Protein; 235 AA.		
XX	AAU27666:		
AC	AAU27666:		
XX	18-DEC-2001 (first entry)		
DT	Human protein AFB0526.		
DE	Human; AFR; cyclostatic; Marfan's syndrome; thrombocytopoenia;		
XX	leukemia; porphyria; Gilles De La Tourette's syndrome; immunogen;		
KM	equamous cell carcinoma; diabetes mellitus; Grave's disease;		
KM	colon cancer; Alzheimer's disease; epiphyseal dysplasia.		
XX	Homo sapiens.		
OS	MO200166748-A2.		
PN	13-SEP-2001.		
XX	05-MAR-2001; 2001MO-US07192.		
XX	03-MAR-2000; 2000US-187221P.		
XX	(ZYMO) ZYMOGENETICS INC.		
PA	Conklin DC, Presnell SR, Adler DA;		
PI	WPI; 2001-589943/66.		
XX	N-PSDB; AAS44567.		
DR			
XX			

Novel AFP polypeptides and polymnucleotides, useful for diagnostic and therapeutic purposes, in cancer therapy and for screening modulator compounds -

Claim 2; Page 194; 220pp; English.

The invention relates to novel human AFP proteins (not defined) and the nucleic acids that encode them. AFP proteins are useful as standards in assays of protein and protein inhibitors in both clinical and research settings, as protein and amino acid supplements, including hydrolyzates. The nucleic acids are useful for radiation hybrid mapping. Secretory fusion proteins of AFP are useful in cancer therapy, for enhancing in vitro cytotoxicity, for enhancing in vivo killing of target tissues, and for targeted cells or tissue inhibition or ablation. Anti-AFP antibodies are useful for isolating target polypeptides by affinity purification, in diagnostic assays for determining circulating or localised levels of target polypeptides, for tissue typing, for cell sorting, for screening expression libraries, for generating anti-idiotypic antibodies, and as neutralising antibodies or as antagonists to block protein activity in vitro and in vivo. AFP proteins and nucleic acids may be used to diagnose or treat (e.g. by gene therapy) diseases associated with the malfunction of the AFP e.g. AFP166924 and Marfan's syndrome, AFP576853/AFP39158 and thrombocytopanias, leukemia, porphyria, Gilles De La Tourette's syndrome, AFP65829 and squamous cell carcinoma, diabetes mellitus, Grave's disease, AFP664311 and colon cancer, AFP25034 and Alzheimer's disease and AFP686580 and epiphyseal dysplasia. Many more examples of disease are given in the specification. The present sequence represents an AFP of the invention.

Alignment Scores:			
Pred. No.:	1..12e-125	Length:	235
Score:	1247.00	Matches:	234
Percent Similarity:	99.57%	Conservative:	0
Best Local Similarity:	99.57%	Mismatches:	1
Query Match:	79.94%	Indels:	0
DB:	22	Gaps:	0
US-10-037-270-482 (1-866) x AAU27666 (1-235)			
QY	42 ATGAAATATGCTCTTCAATTTGGGTGTCCTCCCTGGGACATTTTCTTGTGACTATCT		101
Db	1 MetLySTyValPheTyrlEuGlYalLeuAlaGlyTrhPhePhePheAlaAspSer		20
QY	102 GTTCAGAAAGAGACCCCTGCCCTCCATTTGGTGTACCTCAAGCTTCAACCCCTGT		161
Db	21 ValGlnLySGlnAspProAlaProTyrlLeuValTyrlLeuLySerHisAspAsnProCys		40
QY	162 GTGGGCGTCTCATAAACCCAGCTGGGTGGTGGCCCGACGCTCACTGCTATTACCAAT		221
Db	41 ValGlyValLeuIleLyPProSerTrpValLeuAlaProAlaHisCySTyrlLeuProAsn		60
QY	222 CTGAAAGTATGCTGGGAAATTTCAAGACAGAGCTCAGAGACGGTACTGAACAGACAATT		281
Db	61 LeuLySValMetLeuGlYAsnPhelySerZrgValAlaGpaGlyThrGlnGlnThrIle		80
QY	282 AACCCCATTCAGATGTCGCCGCTACAGGAACTACAGTATAGCGGCCACAGATGATCTC		341
Db	81 AsnProIleGlnIleValAlaTrgTyrlTrpAsnTyrlSerHisSerAlaProGlnAspAspLeu		100
QY	342 ATGCTCATCAAGCTGGCTTAAGCTTCACATGCTCAATCCCAAGTCACAGCCCTCCCTC		401
Db	101 MetLeuIleLyLeuAlaLyPProAlaMetLeuAsnProLySValGlnProLeuThrLeu		120
QY	402 GCCACCAACCAATGTCAAGCCAGGACCTGTCTGTCTACTCTCAGTTTGAAGTGAACCA		461
Db	121 AlaThrTrhAsnValAlaTrpGlyTrhValYAspLeuLeuSerGlyLeuAspTrpSerGln		140
QY	462 GAAACAGTGGCGGACACCCCTGAAGCTTGGGGAGAAACCTGAGAGGCCCGCGATGTCGAT		521
Db	141 GlnAsnSerGlyAlaHisPProAspLeuAspGlnAsnLeuGlnAlaProValMetSerAsp		160

PI Rastelli L, MacDougall JR, Taupier RJ, Grosse WM, Szekeres ES;
PI Alsbrook JP;
XX
XX
DR WPI: 2002-049278/06.
DR N-PSDB; ABA04647.
XX
XX
PT Novel G-protein coupled receptor-related polypeptides and
PT polynucleotides for diagnosing, preventing and treating cardiomyopathy,
PT atherosclerosis, disorders related to cell signal processing and for
PT identifying modulators -
XX
XX
PS Example 1; Page 216; 227pp; English.

CC The present invention relates to novel G-coupled protein-receptor related
CC proteins and coding sequences (MOLX, where X is a number from 1 to 10,
CC ABA04589-ABA04603 and AAM47659-AAM47673). MOLX proteins and coding
CC sequences are useful for treating or preventing a MOLX-associated
CC disorder, such as cardiomyopathy, atherosclerosis, disorders related
CC to cell signal processing and metabolic pathway modulation, diabetes and
CC cancer. Additionally, MOLX proteins and coding sequences are useful for
CC preventing and treating a variety of disorders including metabolic
CC disorders, nutritional oedema, chronic and hereditary pancreatitis,
CC obesity, infectious disease, anorexia, neurodegenerative disorders,
CC Alzheimer's disease, Parkinson's disease, stroke, immune disorders,
CC haematopoietic disorders and various dyslipidaemias, metabolic syndrome X
CC and wasting disorders associated with chronic diseases and cancers,
CC cardiac disorders, hypertension, hypercalcaemia, cirrhosis, angiogenesis
CC and wound healing, trauma, glomerulonephritis, hyper and hypothyroidism,
CC multiple sclerosis, lung diseases including asthma, Crohn's disease,
CC scleroderma, autoimmune diseases, developmental disorders and neural tube
CC defects. The present sequence is the protein sequence for a variant of
CC MOLA. MOLA is a trypsin-like protein.
XX
XX

SO Sequence 228 AA:

Alignment Scores:

Pred. No.: 2,46e-122 Length: 228
Score: 1216.00 Matches: 226
Percent Similarity: 99.12% Conservative: 0
Best Local Similarity: 99.12% Mismatches: 2
Query Match: 77.95% Indels: 1
DB: Gaps: 0

US-10-037-270-482 (1-866) x AAM47674 (1-228)

QY 42 ATGAATATGCTCTTCTTATTTGGGTGCTGCTGGGACATTTTCTTGTGCTCATCT 101
DB 1 MetLysTyrValPheTyrLeuGlyValLeuAlaGlyThrPhePheAlaSerSer 20
QY 102 GTTCAGAAAGAGACCTGCTGCTCTTGGTGTACCTCAAGCTCACTTCAACCCCTGT 161
DB 21 ValGlnLysGlnLysProAlaProTyrLeuValTyrLeuLysSerHisPheAnProCys 40
QY 162 GTGGGGCTCTCATCAACCCAGCTGGTGGCTGGCCCACTCACTGCTATTACCAAT 221
DB 41 ValGlyValLeuLeuLeuLysProSerTrpValLeuAlaProAlaHisCysTyrLeuProAn 60
QY 222 CTGAAGATGATGCTGGGAAATTTCAAGAGACAGAGCTGACTGTAACGACCAAT 281
DB 61 LeuLysValMetLeuLeuLysAnPheLysSerArgValArgSerPGLYthrgLugIntHrIle 80
QY 282 AACCCCATTCAGATGCTGGCTGCTGGAATCAAGCATAGAGCCGCCCAAGAGATGACCTC 341
DB 81 AspProLysGlnLysValArgTyrTrpAsnTyrSerHisSerLysPheAnProGlnSerPheLeu 100
QY 342 ATGCTCATCAAGCTGCTAAGCTGCTGCTGCTCAATCCCAAGTCCAGCCCTTCCCTC 401
DB 101 MetLeuLysLeuLysLeuLysProAlaMetLeuAnProLysValGlnProLeuThrLeu 120
QY 402 GCCACCAACCAATGTCAGGCCGACCTGTGTCTACTCTGAGTTTGAAGAGCCCA 461
DB 121 AlaThrThrAsnValArgProGlyThrValCysLeuLeuSerGlyLeuAnSerTrpSerGln 140

QY 462 GAAAACAGTGGCCGACACCTGACTTGGCGGAGAACTTGAGAGCCCGCTGATGTGAT 521
DB 141 GUAAnSerGlyArgHisProAspLeuArgGlnAnLeuGlnAlaProValMetSerArg 160
QY 522 CGAGAAATGCCA-AAAAACGAAACAAGAAAAAGCCACAGAAATTCCTTATGTGTGAAT 580
DB 161 ArgGlnCysGlnLysAnArgThrArgLysLysProGlnGlnPheLeuMetCysGlnIle 180
QY 581 TGTGAAAGATTCAGCCGAATTTTGGGAGGTGGCCGCTGCTACTGTCATGCAAGA 640
DB 181 CysGlnSerLysGlnPheAnSerPheTrpGlyGlyArgCysTyrCysHisLeuGlnArg 200
QY 641 CAAGCTCCAGGAAATCGAGTGGGCACTTCATGGAGGAGGAGCTGCGCATTCACCA 700
DB 201 GlnAlaProGlyAsnArgGlyGlyAlaLeuHisGlyArgGlyArgArgHisLeuHisGln 220
QY 701 TGTTTACAAATATGATTCCTGAT 724
DB 221 CysLeuGlnHisCysLeuLeuAnp 228

RESULT 12

AAB21321
ID AAB21321 standard; Protein, 247 AA.

AC AAB21321;

DT 02-FEB-2001 (first entry)

XX Human trypsinogen.

KM Human; KLK-L1; KLK-L2; KLK-L3; KLK-L4; KLK-L5; KLK-L6; trypsinogen;

KM kallikrein-like protein; serine protease;

XX cytostatic; cancer; prostate cancer.

OS Homo sapiens.

PN WO200053776-A2.

PD 14-SEP-2000.

PF 09-MAR-2000; 2000MO-CA00258.

PR 11-MAR-1999; 99US-0124260.

PR 01-APR-1999; 99US-0127386.

PR 21-UTL-1999; 99US-0144919.

PA (MOUN) MOUNT SINAI HOSPITAL.

PI Yousef GM, Diamandis EP;

DR WPI: 2000-587440/55.

XX New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L

PT protein mediated disorders, especially cancer. -

PS Example 4; Fig 17; 184pp; English.

XX The present sequence is human trypsinogen, a member of the serine

CC protease family. Kallikreins and kallikrein-like proteins are a

CC subgroup of the serine protease enzyme family. They catalyse the

CC selective cleavage of specific polypeptide precursors to release

CC peptides with potent biological activity. Nucleic acids encoding

CC kallikrein-like proteins KLK-L1, KLK-L2, KLK-L3, KLK-L4, KLK-L5 and

CC KLK-L6 have been isolated. The proteins are useful in the treatment,

CC monitoring and diagnosis of cancers, especially prostate cancer. They

CC can also be used to identify a substance that can associate with or

CC mediate the biological activity of the proteins. Antibodies can

CC be used to treat conditions mediated by the kallikrein-like proteins.

SO Sequence 247 AA;

Alignment Scores:

Pred. No.: 4.93e-29 Length: 247

Score: 352.50 Matches: 82
 Percent Similarity: 50.43% Conservativity: 36
 Best Local Similarity: 35.04% Mismatches: 109
 Query Match: 22.60% Indels: 7
 DB: 21 Gaps: 2

US-10-037-270-482 (1-866) x AAY78975 (1-247)

QY 54 TTCTATTGGGTGCTCTGCTGGAGACATTTTCTTCTGACATCTGTCTAG----- 107
 Db 9 PheValAlaAlaLeuAlaLeuAlaProPheAspAspAspSerLysIleValGlyGlyTyr 28
 QY 108 -----AAGAGAGACCTGCTCTCTATTTGGTGTACTCAAGTCTCACTTCAACCTCT 161
 Db 29 AsnGlySerIleuAsnSerValProTyrGlnValSerIleuAsnSerGlyTyrHisAspGly 48
 QY 162 GTGGGCGCTCTCATCAACCCAGCTGGGTGCTGGCCCGACGCTCACTGTTTATCAAAAT 221
 Db 49 GlyGlySerIleuIleAsnGlnGlnTTPValValSerIleuGlyHisCysTyrLysSerArg 68
 QY 222 CTGAAAGTGTGCTGGGAAATTTTCAGAGCAGAGTCAAGCGGTACTGAAACAGACAAAT 281
 Db 69 ILeGlnValArgLeuGlyGlnHisAsnIleGlnValLeuGlnGlyAsnGlnPheIle 88
 QY 282 AACCCATTCAAGTCTGCTCTGCTGAACTCAAGTCAAGCGCCCAAGATGACCTC 341
 Db 89 AsnAlaIleAlaIleIleArgHisProGlnTyrAspArgLysThrLeuAsnAspIle 108
 QY 342 ATGCTCATCAAGCTGCTAGCCGCTGCAATCCCAAGTCAAGCGCCCTCCCTC 401
 Db 109 MetIleuIleuSerSerSerArgAlaValIleAsnAlaArgValSerThrIleSerLeu 128
 QY 402 GCCAACCCATGTCAAGCCAGCAGCAGTCTGTCTACTCTCAAGTTCAGTGGAGCAA 461
 Db 129 ProThrAlaProProAlaThrGlnThrLysCysIleuIleSerGlyTyrGlnAsnThrAla 148
 QY 462 GAAACAGTGGCCGACACCTGACTTGGCAGAACTGAGGCCCCGTGATGCTGAT 521
 Db 149 SerSerGlyAlaAspTyrProAspGlnLeuGlnCysIleuAspAlaProValIleuSerGln 168
 QY 522 CGAAGATGCGCAAAAGAGAAAGAAAGCAAGAAAGCAAGAAATTCATATGAGAAAT 581
 Db 169 AlaIleCysGlnAlaSerTyrProGlyLysIleThrSerAsnMetPheCysValGlyPhe 188
 QY 582 GTGAAATATTCAGCCGAATTTTGGAGAGTGGCCGTGCTACTGATCTGCAAGAC 641
 Db 189 LeuGlnGlyLysAspSerCysGlnGlyAspSerGlyProValValCysAsnGly 208
 QY 642 AAGCTCAGGGAATCCAGGTG-----GGCATTCTCATGGAGGGAGCGTGGCATC 692
 Db 209 GlnLeuGlnGlnIleValIleSerTyrGlyAspGlyCysAlaGlnLysAsnLysProGlyVal 228
 QY 693 TACACCAATGTTTACAAATATGATCTGATGAGAACT 734
 Db 229 TyrThrLysValTyrAsnTyrValLysTyrIleLysAsnThr 242

RESULT 13
 ID AAY78975 standard; protein; 246 AA.
 AC AAY78975;
 XX
 DT 05-JUN-2000 (first entry)
 XX
 DE Canine anionic trypsinogen amino acid sequence.
 XX
 KW Anionic trypsinogen; dog; monoclonal antibody production; detect;
 KM diagnosis; pancreatitis; pancreatic cancer; renal insufficiency;
 XX extrapancreatic hyposecretion.
 OS
 XX Canis familiaris.
 PN MO200009739-A1.

XX 24-FEB-2000.
 PD
 XX 09-AUG-1999; 99MO-JP04299.
 PF
 XX 10-AUG-1998; 98UP-0236609.
 PR 10-MAR-1999; 99UP-0063990.
 XX
 XX (FUJY) FUJI YAKUHIN KOGYO KK.
 PA
 PI Waricani T, Ashida Y, Yamada T;
 XX
 XX WPI; 2000-206018/18.
 DR
 XX
 XX Anti-canine trypsin monoclonal antibody, useful as reagent for quick
 PT and accurate detection and quantitation of trypsin and/or trypsin-like
 PT immunoreactants in various forms in diagnosis e.g. of pancreatic
 PT diseases -
 XX
 XX Claim 3; Page 64-65; 67pp; Japanese.
 PS
 XX
 CC This sequence represents the canine anionic trypsinogen amino acid
 CC sequence. The protein was isolated from the canine pancreas. The
 CC invention relates to monoclonal antibodies with specificity against
 CC canine trypsin, or canine trypsin-related substances. The antibodies are
 CC highly specific and can be used as reagent for quick and accurate
 CC detection and quantitation of canine trypsin and canine trypsin-like
 CC immunoreactants in various forms. The antibodies can be used in the
 CC diagnosis of diseases such as pancreatitis, pancreatic cancer, renal
 CC insufficiency and extrapancreatic hyposecretion.
 CC
 SQ Sequence 246 AA;

Alignment Scores:
 Pred. No.: 8,09e-29 Length: 246
 Score: 350.50 Matches: 81
 Percent Similarity: 50.83% Conservativity: 42
 Best Local Similarity: 33.47% Mismatches: 108
 Query Match: 22.47% Indels: 11
 DB: 21 Gaps: 2

US-10-037-270-482 (1-866) x AAY78975 (1-246)

QY 42 ATGAAATATGCTCTTATTGGGTGCTCTGCTGGAGACATTTTCTTCTGACATCT 101
 Db 1 MetLysThrPheIlePheLeuAlaLeuGlnValAlaThrValAlaPheProIleAspAsp 20
 QY 102 GTTCAGAA-----GAAAGCCGTGCTCTTATTGGGTGAC 137
 Db 21 AspAspLysIleValGlyGlyTyrThrCysSerArgAsnSerValProTyrGlnValSer 40
 QY 138 GTCAAGTCACTGCAACCCGTGCGGCTGCTCATGAACCCAGCGGGTGGCC 197
 Db 41 LeuAsnSerGlyTyrHisPheCysGlyLysSerIleAsnSerGlnTyrValValSer 60
 QY 198 CCAGCTCACTGCTATTTCACCAATCTGAAGTATCTGGAAATTTCAAGACAGAGTC 257
 Db 61 AlaAlaHisCysTyrLysSerArgIleGlnValArgLeuGlyGlyTyrAsnIleAlaVal 80
 QY 258 AGAGAGCTAGTGAACAAATTAACCCATTCAAGTGTCCGCTACTGGAATCACT 317
 Db 81 SerGlnGlyGlyGlnGlnPheIleAsnAlaAlaLysIleIleArgHisProArgTyrAsn 100
 QY 318 CATAGCGCCCAAGAGATGACTCAAGTCACTCAAGTGGCTAAGCTCCATGCTCAT 377
 Db 101 AlaAsnThrIleAspAsnAspIleMetLysIleuSerSerProAlaThrLeuAsn 120
 QY 378 CCAAGTCCAGCCCTTCCCTGCGCACCAATGTCAGGCCAGGACATGCTGTCTTA 437
 Db 121 SerArgValSerAlaIleAlaLeuProLysSerCysProAlaAlaGlyThrGlnCysLeu 140
 QY 438 CTCTCAGTTTGGACTGAGCCAGAAACAGTGGCCGACACCTGATCTTGGCGAGAAC 497

Db 141 IleserGlyTrrpGlyAsnThrGlnSerIleGlyGlnAsnTyrProAspValLeuGlnCys 160
 QY 498 CTGGAGCCCGCCGGTGTCTGATCGAAGATGCCAAAAACAGACAGAGAAAACCCAC 557
 Db 161 LeuValAspProIleLeuSerAspSerValCysArgAsnAlaTyrProGlyGlnIleSer 180
 QY 558 AGGAATTCCTTATGTGTGAAATTTGTGAAATTTATGACCGAATTTTGGGAGGTGGC 617
 Db 181 SerAsnMetMetCysLeuGlyTyrMetGlnGlyGlyLeuAspSerCysGlnGlyAspSer 200
 QY 618 GTTGCTACTGTCTCTGCAAGACAGACTCCAGGGAATC-----GAGGTGGGAC 668
 Db 201 GlyGlyProValValCysAsnGlyGlnLeuGlnGlyValValSerTrrpGlyAlaGlyCys 220
 QY 669 TTCATGGAGGAGGAGCTGGCATCTACCAATGTTTAAATATGATATTCCTGATTTAG 728
 Db 221 AlaGlnLeuGlyLeuAspProGlyValSerProValValCysLeuTyrValSerTrrpIleGln 240
 QY 729 AACACT 734
 Db 241 GlnThr 242
 RESULT 14
 AAY31160
 ID AAY31160 standard; protein; 224 AA.
 XX AAY31160;
 AC
 DT 26-OCT-1999 (first entry)
 XX
 DE Human trypsin serine protease protein domain.
 XX
 KM Macrophage stimulating protein; MSP; human; modulator; proliferation;
 KW differentiation; intestinal epithelium; colon crypt; treatment; cancer;
 KM haematopoietic disorder; megakaryocyte deficiency; gastrointestinal;
 KM chemotherapeutic agent; gut toxicity; serine protease; trypsin.
 XX
 OS Homo sapiens.
 XX
 FN US594892-A.
 PD 07-SEP-1999.
 XX
 PF 16-DEC-1996; 96US-0766982.
 XX
 PR 16-DEC-1996; 96US-0766982.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Wahl RC;
 XX
 DR WPI, 1999-517975/43.
 XX
 PT Analogues of macrophage stimulating protein for treating
 XX gastrointestinal or haematopoietic disorders
 PS Example 2; Column 27-30; 23pp; English.
 XX
 CC This invention describes a novel purified and isolated analogue of mature
 CC macrophage stimulating protein (MSP) having at least one unpaired
 CC cysteine residue substituted with another amino acid which modulates the
 CC proliferation or differentiation of the intestinal epithelium. The
 CC product of the invention binds to RON (a cell membrane protein tyrosine
 CC kinase which is a member of the c-met family) to promote the formation of
 CC colon crypts. MSP analogues are useful for the treatment of conditions
 CC requiring the administration of MSP, such conditions include
 CC haematopoietic disorders such as those involving a deficiency of
 CC megakaryocytes and gastrointestinal disorders such as ulcerative colitis,
 CC Crohn's disease and infections. The MSP analogues are useful for
 CC maintaining and repairing the epithelial lining in the treatment of
 CC cancer, where the aggressive use of chemotherapeutic agents or the use of
 CC whole body radiation may lead to gut toxicity. The MSP analogues, which
 CC have a higher activity than normal human MSP are effective at smaller

CC dosages, or optionally, they may be administered less frequently than
 CC human MSP. This sequence represents a human trypsin serine protease
 CC domain which is used in a description of the method of the invention.
 XX
 SQ Sequence 224 AA;
 Alignment Scores:
 Pred. No.: 1e-28 Length: 224
 Score: 349.50 Matches: 76
 Percent Similarity: 52.83% Conservative: 35
 Best Local Similarity: 35.85% Mismatches: 97
 Query Match: 22.40% Indels: 3
 DB: Gaps: 1
 US-10-037-270-482 (1-866) x AAY31160 (1-224)
 QY 108 AAGAAGACCCCTGCTCCCTATTTGGTGAATCTCAATCTCAATCAACCCCTGTGGC 167
 Db 8 GlnGlnAsnSerValProTyrGlnValSerLeuAsnSerGlyTyrHisPheCysGlyGly 27
 QY 168 GTCCCATCAAAACCCAGCTGGGTGGTGGCCGAGCTCACTGATTTACCAATCGAAA 227
 Db 28 SerLeuIleAsnGlnGlnTrrpValValSerAlaGlyHisCysTyrIleSerArgIleGln 47
 QY 228 GTGATGCTGGAAATTTTCAAGACAGAGTACAGACGCTACTGACAGACAAATTAACCC 287
 Db 48 ValArgLeuGlyGlnHisAsnIleGlnValLeuGlnGlyAsnGlnIleHisAla 67
 QY 288 ATTCAAGATGCTCCGCTACTGGAATCACTCACTATAGCCGCCACAGATGACTCATGCTC 347
 Db 68 AlaValIleIleAlaArgHisProGlnTyrAspArgIleTyrLeuAsnAsnAspIleMetLeu 87
 QY 348 ATCAAGCTGGCTAAGCTGCTCCATGCTCAATCCCAAGTCCAGCCCTTCCCTGGCCACC 407
 Db 88 IleValLeuSerSerArgAlaValIleAsnAlaArgValSerThrIleSerLeuProThr 107
 QY 408 ACCAATGTACGACCCAGCTGCTGTCTACTCTCACTGCTTGGATGAGCCCAAGAAAC 467
 Db 108 AlaProProAlaThrGlyThrIleCysLeuIleSerGlyTrrpGlyAsnThrAlaSerSer 127
 QY 468 AGTGGCCACACCTGACTTGGGGAGAACTGGAGGCCCGGTGTGTGTGATCGAAGAA 527
 Db 128 GlyAlaAspTyrProAspIleLeuGlnCysLeuAspAlaProValLeuSerGlnAlaLys 147
 QY 528 TGGCAAAAAACAGACAGAAAGAAAGCCACAGAAATCTCTTATGTGTGAAATTTGAAA 587
 Db 148 CysGlnAlaSerTyrProGlyIleThrSerAsnMetPheCysValAlaIlePheLeuGln 167
 QY 588 GTATTACGCCGAATTTTGGGGAGGTGGCGTGTGCTACTGCTCATCTGCAAGACAGCTC 647
 Db 168 GlyGlyLeuAspSerCysGlnGlyAspSerGlyGlyProValValCysAsnGlnLeu 187
 QY 648 CAGGAATCAGAGTG-----GGGCACTTCATGGAGGGGAGCGTGGCATCTACAC 698
 Db 188 GlnGlyValValSerTrrpGlyAspGlyCysAlaGlnLeuAsnLysProGlyValTyrThr 207
 QY 699 AATGTTTACAAATATGTATCTCGATTCGATTCGAAACACT 734
 Db 208 LysValTyrAsnTyrValLysTrrpIleLysAsnThr 219
 RESULT 15
 AAB98503
 ID AAB98503 standard; Protein; 225 AA.
 XX AAB98503;
 AC
 DT 03-AUG-2001 (first entry)
 XX
 DE Human trypsin serine protease catalytic domain.
 XX
 KM Human; TADG-15; cytostatic; vaccine; ovarian tumour; cancer; trypsin;
 KM tumour antigen-derived gene 15; serine protease.

OS	Homo sapiens.
XX	
PN	MO200129056-A1.
XX	
PD	26-APR-2001.
XX	
PF	20-OCT-2000; 2000WO-US29095.
XX	
PR	20-OCT-1999; 99US-0421213.
XX	
PA	(UYAR-) UNIV ARKANSAS.
XX	
P1	O'Brien TJ, Tanimoto H;
XX	
DR	WPI; 2001-381031/40.
XX	
XX	
PT	Novel extracellular serine protease, termed tumor antigen-derived gene
PT	15 protein overexpressed in carcinomas and DNA encoding it, for
PT	diagnosis, treatment, prevention of cancer, particularly breast,
PT	ovarian cancer -
XX	

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame plus n2p model

Run on: November 8, 2003, 02:06:22 ; Search time 19 Seconds
(without alignments)
3856.969 Million cell updates/sec

Title: US-10-037-270-482
Perfect score: 1560
Sequence: 1 ggcacttaccctgcagccta.....tgggaatgcagcactagct 866

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 328717 seqs, 42310858 residues
Total number of hits satisfying chosen parameters: 657434

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ n2p model -DRV=x1h
-Q=/cgn2_1/USFTO.spool/US10037270/runat_07112003_140516_8822/app_query.fasta_1.1031
-DB=Issued_Patents_AA-QPMT=fastan-SUFFIX=rat-MINMATCH=0.1-LOOPL=0
-LOOPEXT=0-UNITS=bits-START=1-END=1-MATRIX=bloms62-TRANS=human40.cdi
-LIST=45-DOCALLIGN=200-THR SCORE=pct-THR MAX=100-THR MIN=0-ALIGN=15
-MODE=LOCAL-OUTPM=ptc-NORMext=HEAPSIZE=500-MINLEN=0-MAXLEN=2000000000
-USER=US10037270@cgn2_1.18@runat_07112003_140516_8822-NCPU=6-ICPU=3
-NO_MMAP-LANG=OVERLY-NEG SCORES=0-WAIT-DSPELOCK=100-LOGLOG
-DEV_TIMEOUT=120-WARN_TIMEOUT=30-THREADS=1-XGAPOP=10-XGAPEXT=0.5-FGAPOP=6
-FCAPEXT=7-YGAPOP=10-YGAPEXT=0.5-DELOP=6-DELEXT=7

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/2/iaa/58_COMB.pep:*
2: /cgn2_6/prodata/2/iaa/58_COMB.pep:*
3: /cgn2_6/prodata/2/iaa/68_COMB.pep:*
4: /cgn2_6/prodata/2/iaa/68_COMB.pep:*
5: /cgn2_6/prodata/2/iaa/pctus_COMB.pep:*
6: /cgn2_6/prodata/2/iaa/backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1247	79.9	235	4	US-09-796-110-2
2	349.5	22.4	224	2	US-08-766-982-13
3	349.5	22.4	224	3	US-08-944-483-36
4	349.5	22.4	224	3	US-09-296-219-13
5	349.5	22.4	225	2	US-09-027-337-5
6	349.5	22.2	225	4	US-09-644-600-5
7	346.5	22.2	246	2	US-08-978-404B-44
8	344.5	22.1	224	3	US-08-944-483-35
9	335.5	21.5	247	2	US-08-956-267A-2
10	334.5	21.4	224	3	US-08-944-483-34
11	319.5	20.5	281	1	US-08-467-155A-7
12	319.5	20.5	281	2	US-08-628-198-7

13	319.5	20.5	281	3	US-09-201-038-7	Sequence 7, Appl1
14	319.5	20.5	281	5	PCT-US96-07343-7	Sequence 7, Appl1
15	312.5	20.0	223	1	US-08-766-091-9	Sequence 9, Appl1
16	312.5	20.0	223	1	US-08-483-859-9	Sequence 9, Appl1
17	312.5	20.0	223	1	US-08-472-173-9	Sequence 9, Appl1
18	312.5	20.0	223	2	US-08-487-167-9	Sequence 9, Appl1
19	312.5	20.0	223	2	US-08-482-816-9	Sequence 9, Appl1
20	312.5	20.0	223	2	US-08-296-149-9	Sequence 9, Appl1
21	312.5	20.0	223	2	US-08-801-499-9	Sequence 9, Appl1
22	312.5	20.0	223	2	US-08-615-271-9	Sequence 9, Appl1
23	312.5	20.0	223	3	US-09-074-660-9	Sequence 9, Appl1
24	312.5	20.0	223	3	US-09-074-659-9	Sequence 9, Appl1
25	312.5	20.0	223	3	US-09-106-468-9	Sequence 9, Appl1
26	312.5	20.0	223	3	US-09-106-468-9	Sequence 9, Appl1
27	312.5	20.0	223	3	US-09-106-467-9	Sequence 9, Appl1
28	312.5	20.0	229	3	US-09-120-582-2	Sequence 2, Appl1
29	304.5	19.5	299	1	US-08-467-155A-8	Sequence 8, Appl1
30	304.5	19.5	299	2	US-08-628-198-8	Sequence 8, Appl1
31	304.5	19.5	299	3	US-09-201-038-8	Sequence 8, Appl1
32	304.5	19.5	299	5	PCT-US96-07343-8	Sequence 8, Appl1
33	287.5	18.4	244	1	US-08-361-395-1	Sequence 1, Appl1
34	280	17.9	286	1	US-08-467-155A-9	Sequence 9, Appl1
35	280	17.9	286	2	US-08-628-198-9	Sequence 9, Appl1
36	280	17.9	286	3	US-09-201-038-9	Sequence 9, Appl1
37	280	17.9	286	5	PCT-US96-07343-9	Sequence 9, Appl1
38	261	16.7	156	3	US-09-261-416-6	Sequence 6, Appl1
39	252	16.2	260	3	US-09-008-271A-7	Sequence 7, Appl1
40	251	16.1	260	3	US-09-070-526-2	Sequence 2, Appl1
41	250	16.0	288	4	US-09-386-642-13	Sequence 13, Appl1
42	249	16.0	271	1	US-08-467-155A-10	Sequence 10, Appl1
43	249	16.0	271	2	US-08-628-198-10	Sequence 10, Appl1
44	249	16.0	271	3	US-09-201-038-10	Sequence 10, Appl1
45	249	16.0	271	5	PCT-US96-07343-10	Sequence 10, Appl1

ALIGNMENTS

RESULT 1
US-09-796-110-2
; Sequence 2, Application US/09796110
; Patent No. 6468776
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; TITLE OF INVENTION: Human Serine Protease
; FILE REFERENCE: 00-16
; CURRENT APPLICATION NUMBER: US/09/796,110
; CURRENT FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-110-2

Alignment Scores:

Pred. No.: 4,236-132
Score: 1247.00
Percent Similarity: 99.57%
Best Local Similarity: 99.57%
Query Match: 79.94%
DB: 4
Length: 235
Matches: 234
Conservative: 0
Mismatch: 1
Indels: 0
Gaps: 0

US-10-037-270-482 (1-866) x US-09-796-110-2 (1-235)

QY	42	ATGAATATGCTCTTCAATTGGTGGTCCGCGGAGACATTTTCTTGCTGACTCATCT	101
DB	1	MetuLrYrValPhetYrLeuSlYalLeuAlaGlyTmPhenPheNlaAspSerSer	20
QY	102	GTTCCAGAAAGAACCCCTGCTCCCTAATTGGTGTAACCTGCTCACTTCAACCCCTGT	161
DB	21	ValGlnLysGluAspProHlaProYrLeuValYrLeuLysSerHisPheAsnProCys	40

```

Oy      162 GTGGGCGTCTCATCAAAACCCAGCTGGGGCTGGCCCCAGCTCACTGCTATTTACCAAT 221
Db      41  ValGlyValLeuIleLeuProSerTrpValLeuAlaProAlaIscyTrpLeuProAsn 60
Oy      222 CTGAAGTGCATCTGGGAAATTTCAAGACAGAGTCAGAGACCGTACTGAACAGACAATT 281
Db      61  LeuIyValMetLeuGlyAsnPhenIySerArgValAlaAspGlyIlnTrgIlnThrIle 80
Oy      342 ATGTCATCAAGCTGGCTTAAGCCCTGCAGCTCAATCCCAAGCTCAAGCCCTCCCTC 401
Db      101 MetLeuIleLeuLeuAlaIysProAlaMetLeuAsnProIyValGlnProLeuThrLeu 120
Oy      402 GCCACCAACCAAGTCAAGCGAGCACTGTCTCTACTCTCAGATTGGACTGAAGCAA 461
Db      121 AlaThrThrAsnValAlaArgProGlyIlnThrValCylLeuLeuSerIlyLeuAspTrpSerGln 140
Oy      462 GAAAACAGTGGCCGACACCCCTGACTTGGCGAGACCTTGAGGCCCCCGTAGTGTGAT 521
Db      141 GluAsnSerGlyAlaArgIleProAspLeuAlaGlnAsnLeuGlyAlaProValMetSerAsp 160
Oy      522 CGAGAAATGCCAAAACCAAGAAACCAAGAAATTCCTTAATGTGAAATTT 581
Db      161 ArgGlyuCysGlnIlyThrIlnGlnIlyIySerHisAlaGlnSerIleuCyValIlySphe 180
Oy      582 GTGAAGATTCAGCCGGAATTTTGGGAGAGTGGCGTTGCTACTGATCTCACTCAAGAC 641
Db      181 ValIyValAlaPheSerArgIlePheGlyGlyValAlaValAlaThrValIleCyIyAsp 200
Oy      642 AAGCTCCAGGAATCGAGGTGGGGCACTTCATGGAAGGAGCGCTGGCATCTACACCAAT 701
Db      201 IyLeuLeuGlnGlyIleGlyValAlaGlyHisPheMetGlyIyAspValGlyIleIyTrpAsn 220
Oy      702 GTTTCACAAATATGATCTCTGATGGAACACCTGCTAAGACAAG 746
Db      221 ValIyTrpIyValSerTrpIleGluAsnThrAlaIyAspIyS 235

RESULT 2
US-08-766-982-13
/ Sequence 13, Application US/08766982
/ Patent No. 594892
/ GENERAL INFORMATION:
/ APPLICANT: Mahl, Robert C.
/ TITLE OF INVENTION: Analogs of Macrophage Stimulating
/ TITLE OF INVENTION: Protein
/ NUMBER OF SEQUENCES: 13
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Amgen Inc.
/ STREET: 1840 Dehavilland Drive
/ CITY: Thousand Oaks
/ STATE: California
/ COUNTRY: USA
/ ZIP: 91320-1789
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/766, 982
/ FILING DATE:
/ CLASSIFICATION: 536
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Winter, Robert B.
/ REFERENCE/DOCKET NUMBER: A-441
/ INFORMATION FOR SEQ ID NO: 13:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 224 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single

```

```

:      TOPOLOGY: linear
:      MOLECULE TYPE: protein
:      US-08-766-982-13

Alignment Scores:
Pred. NO.:      8,44e-31      Length:      224
Score:      349.50      Matches:      76
Percent Similarity:      52.83%      Conservative:      36
Best Local Similarity:      35.85%      Mismatches:      97
Query Match:      22.40%      Indels:      3
DB:      2      Gaps:      1

US-10-037-270-482 (1-866) x US-08-766-982-13 (1-224)

QY      108 AAMGAAGACCTGCTCCCTATTTGGTGTACTTCAAGTCTCACTTCAACCCCTGTGTGGC      167
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      8 GTCCTATCAACCCAGCTGGGCGTGGGCCCAAGCTCACTGCTATTTACCAATGCGAA      227

QY      168 GTCCTATCAACCCAGCTGGGCGTGGGCCCAAGCTCACTGCTATTTACCAATGCGAA      227
      |||||
DB      28 SerLeuIleasnGluIntrValValSerAlaGlyHisCysTyrIleSerArgIleGln      47

QY      228 GTGATGCTGGGAATTTTCAAGACAGACAGACAGACAGCTGATGCAAGCAATTAACCC      287
      |||||
DB      48 ValArgLeuIleGluHisIleAsnIleGluValIleGluValAsnGluInIleIleAsnIla      67

QY      288 ATTCAAGTGTCCGCTACTGGAACCTACAGTCAATAGCGCCCAAGAGATGACTATGCTC      347
      ::::::::::::::::::::::
DB      68 AlaValIleIleArgHisProGlnTyrAspArgLysThrLeuAsnAsnAspIleMetLeu      87

QY      348 ATCAAGCTGGCTAAGCTCGCATGCTCATATCCCAAGTCAAGCTCCCTCCCGCCACC      407
      |||||
DB      88 IleValLeuSerSerArgAlaValIleAsnAlaArgValSerThrIleSerLeuProThr      107

QY      408 ACCAATGCTCAGGCAGGACGCTGTCTGTACTCTCAAGTTTGAAGTGAAGCCAGAAAAC      467
      |||||
DB      108 AlaProAlaIleArgIleThrIleCysGluIleIleSerGlyTyrProGlyAsnThrIleSerSer      127

QY      468 AGTGGCCGACACCTTGACTTGGCGGCAACCTGAGAGCCCTCGTAGTGTGATGCGACA      527
      ::::::::::::::
DB      128 GlyAlaLeuArgProAspGluIleGlnCysLeuAspAlaProValLeuSerIleAlaLys      147

QY      528 TGGCAAAAACAGAACAGAAAACCCACAGAAATTCCTTATGATGTAATTTGTGAAA      587
      |||||
DB      148 CysGluIleAsnSerIleProGlyLysIleThrSerAsnMetPheCysValGlyPheLeuGlu      167

QY      588 GTATTCAGCCGAATTTTGGGGAGTGGCGGTGCTACTGTCATGTCAGCAAAAGACGTC      647
      ::::::::::::::
DB      168 GlyValLysAspSerCysGlnGlyLysAspSerGlyProValValCysAsnGlyLeuIleu      187

QY      648 CAGGAATCGAGGTG-----GGGCACTTCATGGAAGGGAGCGTCGGCATCTTACCC      698
      |||||
DB      188 GlnGlyValValSerIleProGlyAspGlyCysAlaGlnLysAsnLysProGlyValTyrThr      207

QY      699 AATGTTTCAAAATATGTATCTCTGGATGTGAGAACACT      734
      |||||
DB      208 LysValTyrAsnTyrValLysTyrIleLysAsnThr      219

RESULT 3
US-08-944-483-36
: Sequence 36, Application US/08944483
: Patent No. 6232456
: GENERAL INFORMATION:
: APPLICANT: COHEN, MAURICE
: APPLICANT: COLPITTS, TRACEY L.
: APPLICANT: FRIEDMAN, PAULA N.
: APPLICANT: GRANADOS, EDWARD N.
: APPLICANT: KLAES, MICHAEL R.
: APPLICANT: RUSSELL, JOHN C.
: APPLICANT: STEWART, KENT D.
: APPLICANT: STROUPE, STEVEN D.
: TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES

```

TITLE OF INVENTION: OF THE PROSTATE
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/944,483
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6183.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 224 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 62324566
US-08-944-483-36
Alignment Scores:
Pred. No.: 8,44e-31 Length: 224
Score: 349.50 Matches: 76
Percent Similarity: 52.83% Conservative: 36
Best Local Similarity: 35.85% Mismatches: 97
Query Match: 22.40% Indels: 3
Gaps: 1
US-10-037-270-482 (1-866) x US-08-944-483-36 (1-224)
QY 108 AAGAAGACCCGCTGCTGCTATTGTTGGTACCTCAAGTCTCACTTCAACCCCTGTGGGC 167
DB 8 GUGUUAenseValProtyrGlnValSerLeuanserelyTyxhisphesyglily 27
QY 168 GTCCATCAAAACCCAGCTGGTGTGGCCGCCAGCTCACTGATTACCAATCGAAA 227
DB 28 SerLeuIleasnIuglnItrpValValserAlaGlyHisCyetylyrSerArgIleGln 47
QY 228 GTGATCTCGGAAATTTCAGAGACAGCTCAGAGCGTACGAAACAAATTAACCC 287
DB 48 ValArgLeuGlyGlnHisasnIleGlnValLeuGlnGlyAsnIuglnIlnPheIleAsnAla 67
QY 288 ATTACATGCTGGCGGACGAGTCACTACAGTCATAGAGCCCAAGATGACCTCATGCTC 347
DB 68 AlAlValIleIleArgHisProGlnItyrAspArgIlyThrLeuAsnAsnAspIleMetLeu 87
QY 348 ATCAAGCTGGCTAGGCTGCTCAATCCCAAGTCCAGCCCTTCCCTCGGCACC 407
DB 88 IleIyLeuSerSerArgIleValIleAsnAlaArgValSerThrIleSerLeuProThr 107
QY 408 ACCAATGTACGGCCAGGCACTGTCTGTACTCTCAGGTTTGACGTGAGCCCAAGAAAC 467
DB 108 AlAProProAlaThrGlyThrIlyScySleuIleSerGlyTyrlpIlyAsnThrAlaSerSer 127
QY 468 AGTGGCGACACCCCTGACTTGGCGGAGAACCTGAGAGCCCCCGTGTATGTGATCGAGA 527

DB 128 GlyAlaAspTyProAspGlnIleuGlnCysLeuAspAlaProValLeuSerGlnAlaTyS 147
QY 528 TGGCAAAAAACAGAACAGAAAGAACCCACAGAGATTCTTGTGTGAATTTGGGAA 587
DB 148 CysGlnAlaSerTyProGlyIlySerIleThrSerAsnMetPheCyValGlyPheLeuGln 167
QY 588 GTATTACAGCCGAATTTTGGGAGAGGTGGCCGTGTCTACTGTCATGCAAAAGACGTC 647
DB 168 GlyGlyIlyAspSerCyGlnGlnIlyAspSerGlyGlyProValValCyAsnGlnIlyGlnLeu 187
QY 648 CAGGAAATCGAGTG-----GGGCACTTCATGGAGGAGGAGCGTGGCATTCAGCC 698
DB 188 GlnGlyValAlaSerItyrGlyAspGlyCyAlaGlnIlyAsnIyAsnIyProGlyValTyThr 207
QY 699 AATGTTTACAAATATGTATCTCGATTCGATGAGAACT 734
DB 208 IySValIyTyAsnIyValIySerIlyIleIyAsnThr 219
RESULT 4
US-09-296-219-13
Sequence 13, Application US/09296219
Patent No. 6248560
GENERAL INFORMATION:
APPLICANT: Wahl, Robert C.
TITLE OF INVENTION: Analogs of Macrophage Stimulating
TITLE OF INVENTION: Protein
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: 1840 Dehavenland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/296,219
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-441
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 224 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-296-219-13
Alignment Scores:
Pred. No.: 8,44e-31 Length: 224
Score: 349.50 Matches: 76
Percent Similarity: 52.83% Conservative: 36
Best Local Similarity: 35.85% Mismatches: 97
Query Match: 22.40% Indels: 3
Gaps: 1
US-10-037-270-482 (1-866) x US-09-296-219-13 (1-224)
QY 108 AAGAAGACCCGCTGCTGCTATTGTTGGTACCTCAAGTCTCACTTCAACCCCTGTGGGC 167
DB 8 GUGUUAenseValProtyrGlnValSerLeuanserelyTyxhisphesyglily 27
QY 168 GTCCATCAAAACCCAGCTGGTGTGGCCGCCAGCTCACTGATTACCAATCGAAA 227
DB 28 SerLeuIleasnIuglnItrpValValserAlaGlyHisCyetylyrSerArgIleGln 47

Oy 228 TFGATGCTGGGAAATTTCAAGAGCAGAGCCAGAGCGGTACTTCAACGACATTAACCCC 267
 Db 48 ValaIgluEgLIgLIuHIsanIleGIuValleuGIgLIyasnIugInpHeIIeasAla 67
 Oy 288 ATTGAGATCGTCGGCTACTGGAACATCAGTCATAGGCGCCACACAGATGACCTCATGCTC 347
 Db 68 AlAlYsIIeIIeAlrghIsProGInTYrAspArGlyThrIeUasnsnAspIIeMetIeu 87
 Oy 348 ATCAGAGCTGGCTAGGCGCTGCATGCTCAATGCCAAAGTCAGGCCCTTCCCTGGCCACC 407
 Db 88 IleYsIeUsErSeSrArGlaValIIeasnAlArGValSerThrIleSeSrIeUpProThr 107
 Oy 408 ACCATGTCAGGCCGACGCACTGCTCTGTAACCTGACGTTGGACCTGAGGCCAAGAAAC 467
 Db 108 AlaProProlaIarHnclYlThrIyScYsIeUlleSeSrGlyTyrpIyasnHraIaSeSr 127
 Oy 468 AGTGCCCGCACCCCTGACTTGGCGGCAAGACCTGAGAGGCCCGCTGATGTCATCGAGAA 527
 Db 128 GlYalAlaSPryrProAspGluIeuGInCYsIeasPAlaProValIeUsErGInAlAlYs 147
 Oy 528 TGGCAAAAAACAACCAAGAAAAAACCCACAGGAATTCCTATGCTGGAATTTGTGAAA 587
 Db 148 CysGluAlaSeTyProGlyLysIleThrSeAsnMeScheCYsAlGlyPhleIeuGlu 167
 Oy 568 GTATTCAGCCGAATTTTGGGAGGAGTGCCGCTGTACTGTCAATCTGCAAAAGCAAGCTC 647
 Db 168 GlYglYIyAspSeSrCYsGlnGlyAspSeSrGlyIProValValCYsasnGlyGlnIeu 187
 Oy 648 CAGGGAATCGAGTG-----GGGCACCTCATGGAGGGAGCGCTGGCATCTACACC 698
 Db 188 GlnGlyAlaValaISeSrTyrpIyAspGlyCYsAlaGlnLYsAsnLYsProGlyAlaTYrThr 207
 Oy 699 AATGTTTCAAAATATGTATCTGATCTGATGGAACT 734
 Db 208 LysValTYrAsnTYrValYsTrIpIleYsAsnThr 219

```

RESULT 5
US-09-027-337-5
: Sequence 5, Application US/09027337B
: Patent No. 5972616
: GENERAL INFORMATION:
: APPLICANT: O'Brien, Timothy J.
: APPLICANT: Tanimoto, Hirocoshi
: TITLE OF INVENTION: TAGG-15: An Extracellular Serine Protease Overexpressed in
: TITLE OF INVENTION: Breast and Ovarian Carcinomas
: FILE REFERENCE: D6064
: CURRENT APPLICATION NUMBER: US/09/027,337B
: CURRENT FILING DATE: 1998-02-20
: NUMBER OF SEQ ID NOS: 13
: SEQ ID NO 5
: LENGTH: 225
: TYPE: PRT
: ORGANISM: Unknown
: FEATURE:
: OTHER INFORMATION: Serine protease catalytic domain of trypsin (Try)
: OTHER INFORMATION: homologous to similar domain in TAGG-15
US-09-027-337-5

```

Alignment Scores:	
Pred. No.:	8 46e-1
Score:	349.50
Percent Similarity:	52.83%
Best Local Similarity:	35.85%
Query Match:	22.40%
DB:	
	2
	1
Length:	222
Matches:	76
Conservative:	35
Mismatches:	97
Indels:	3
Gaps:	1

US-10-037-270-482 (1-866) x US-09-027-337-5 (1-225)

Qy 108 A A A G A G A C C C T G C T C C C A T T T T G T G T A C C T C A A G T C T C A C T T C A A C C C C T G T G G G C 167
:::||||: ||||| ||||| ||| :::: |||||
Db 9 G l u c i n a m s e r v a l p r o t y r c i n v a l s e r l e u a m s e r g l y t y r h i s p h e c y s g l y g l y 28

[illegible]

```

RESULT 6
US-09-644-600-5
; Sequence 5, Application US/09644600
; Patent No. 6451500
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: TDMG-15: An Extracellular Serine Protease
; TITLE OF INVENTION: Overexpressed in Carcinomas
; FILE REFERENCE: D6064CIP/D
; CURRENT APPLICATION NUMBER: US/09/644,600
; CURRENT FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/421,213
; PRIOR FILING DATE: 1999-10-20
; PRIOR APPLICATION NUMBER: 09/027,337
; PRIOR FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 5

```

```

/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: Tyrosin
US-09-644-600-5

```

Score: _____

Best Local Similarity:	35.85%	Mismatches:	97
Query Match:	22.40%	Indels:	3
DB:	4	Gaps:	1

US-10-037-270-482 (1-866) x US-09-644-600-5 (1-225)

QY 108 AAGAGACCTGCTCCCTATTGTTGTTAAGCTTCAAGCTTCAACCCCTGTGGGCG 167
 Db 9 GUGUAGAServAlProtyrGlnValSerLeuAsnSerGlyTyrHisPheCysGlyGly 28
 QY 168 GTCTCATGAAACCCAGCTGGGCTGGGCCCCAGCTCACTGCTATTACCAATTCGAAA 227
 Db 29 SerLeuHisAsnGlnIntrpValValSerAlaGlyHisCysTyrIleSerArgIleGln 48
 QY 228 GTGATCTGGGAAATTCAGACGAGCTGAGACGCTGACGAGCTGACAGCAATTAACCC 287
 Db 49 ValArgLeuGlyGlnHisAsnIleGluValLeuGlnGlyAsnGlnGlnPheIleAsnAla 68
 QY 288 ATTACATGCTGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 347
 Db 69 AlaValIleIleLeuArgHisProGlnTyrAspArgIleThrLeuAsnAsnAspIleMetLeu 88
 QY 348 ATCAAGCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 407
 Db 89 IleValLeuSerSerArgAlaValIleAsnAlaArgValSerThrIleSerLeuProThr 108
 QY 408 ACCAATGTACGAGCGACGCTGTCTGTCTGCTGCTGAGCTGAGCTGAGCGCAAGAAAC 467
 Db 109 AlaProProAlaThrGlyThrGlyCysLeuIleSerGlyTyrGlyAsnThrAlaSerSer 128
 QY 468 AGTGGCGGACACCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 527
 Db 129 GlyAlaAspTyrProAspGlnLeuGlnCysLeuAspAlaProValLeuSerGlnAlaLys 148
 QY 528 TCCCAAAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 587
 Db 149 CysGlnAlaSerTyrProGlyLysIleThrSerAsnMetPheCysValGlyPheLeuGln 168
 QY 588 GTATTACGCGAATTTTGGGAGGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 647
 Db 169 GlyGlyLysAspSerCysGlnGlyAspSerGlyGlyProValValCysAsnGlyGlnLeu 188
 QY 648 CAGGAAATGAGGTG-----GGGCACTTATGGAGGAGGAGGAGGAGGAGGAGGAGG 698
 Db 189 GlnGlyValValSerTyrGlyAspGlyCysAlaGlnLysAsnLysAspProGlyValTyrThr 208
 QY 699 AATGTTACAAATATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 734
 Db 209 LysValTyrAsnTyrValLysTyrIleLysAsnThr 220

RESULT 7
 US-08-978-404B-44
 ; Sequence 44, Application US/08978404B
 ; Patent No. 5968782
 ; GENERAL INFORMATION:
 ; APPLICANT: Stevens, Richard L.
 ; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
 ; TITLE OF INVENTION: FIBRINOGEN
 ; NUMBER OF SEQUENCES: 74
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Wolf, Greenfield & Sacke, P.C.
 ; STREET: 600 Atlantic Avenue
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: U.S.A.
 ; ZIP: 02210-2211
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/978,404B
 ; FILING DATE: 25-NOV-97
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/032,354

FILING DATE: 04-DEC-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Plumer, Elizabeth R.
 ; REGISTRATION NUMBER: 36,637
 ; REFERENCE/DOCKET NUMBER: B0801/7090
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-720-3500
 ; TELEFAX: 617-720-2441
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 44:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 246 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: NO. 5968782
 ; US-08-978-404B-44

Alignment Scores:
 Pred. No.: 1,93e-30 Length: 246
 Score: 346.50 Matches: 76
 Percent Similarity: 54.50% Conservative: 39
 Best Local Similarity: 36.02% Mismatches: 93
 Query Match: 22.21% Indels: 3
 DB: 2 Gaps: 1

US-10-037-270-482 (1-866) x US-08-978-404B-44 (1-246)

QY 111 GAAGACCTGCTCCCTATTGTTGTTAAGCTTCAAGCTTCAACCCCTGTGGGCGCT 170
 Db 32 GlnHisSerValProtyrGlnValSerLeuAsnSerGlyTyrHisPheCysGlyGlySer 51
 QY 171 GTCATGAAACCCAGCTGGGCTGGGCCCCAGCTCACTGCTATTACCAATTCGAAAAG 230
 Db 52 LeuIleAsnAspGlnTyrValValSerAlaValHisCysTyrLysSerArgIleGlnVal 71
 QY 231 ATGCTGGGAAATTTCAAGAGCAGCTGAGACGCTGAGCTGAGCTGAGCTGAGCTGAGCT 290
 Db 72 ArgLeuGlyGlnHisAsnIleAsnValLeuGlnGlyAspGlnGlnPheIleAsnAla 91
 QY 291 CAGATGCTCCGCTAGTGAAGCTGACATGACGCTGAGCTGAGCTGAGCTGAGCTGAGCT 350
 Db 92 LysIleIleLeuHisProAsnTyrSerSerTyrThrLeuAsnAsnAspIleMetLeuIle 111
 QY 351 AAGCTGGCTAAGCTGGCTGCTCAATCCCAAGTGCACGCTTCCCTCCGCAACACC 410
 Db 112 LysLeuSerSerProValLysLeuAsnAlaArgValAlaProValAlaLeuProSerAla 131
 QY 411 AATGTCAGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 470
 Db 132 CysAlaProAlaGlyThrGlnCysLeuIleSerGlyTyrGlyAsnThrLeuSerAsnGly 151
 QY 471 GCGCCAGACCTGACTTGGGCGCAAGCTGAGAGGCCCCGATGCTGATGCGAATGC 530
 Db 152 ValAsnAsnProAspLeuLeuGlnCysValAspAlaProValLeuSerGlnAlaAspCys 171
 QY 531 CAAAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 590
 Db 172 GlnAlaAlaTyrProGlyGlnIleThrSerSerMetIleCysValGlyPheLeuGlnGly 191
 QY 591 TTCAGCGGAATTTTGGGAGGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 650
 Db 192 GlyLysAspSerCysGlnGlyAspSerGlyGlyProValValCysAsnGlyGlnLeuGln 211
 QY 651 GGAATC-----GAGGTGGGCACTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 701
 Db 212 GlyIleValSerTyrGlyTyrGlyCysAlaLeuProAsnAsnProGlyValTyrThrLys 231
 QY 702 GTTTACAAATATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 734
 Db 232 ValCysAsnPheValGlyTyrIleGlnAspThr 242

RESULT 8

```

US-08-944-483-35
: Sequence 35, Application US/08944483
: Patent No. 6232456
: GENERAL INFORMATION:
: APPLICANT: COHEN, MAURICE
: APPLICANT: COLPITTS, TRACEY L.
: APPLICANT: FRIEDMAN, PAULA N.
: APPLICANT: GRANADOS, EDWARD N.
: APPLICANT: KLASS, MICHAEL R.
: APPLICANT: RUSSELL, JOHN C.
: APPLICANT: STEWART, KENT D.
: APPLICANT: STROUPE, STEVEN D.
: TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
: TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
: TITLE OF INVENTION: OF THE PROSTATE
: NUMBER OF SEQUENCES: 76
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Abbott Laboratories
: STREET: 100 Abbott Park Road
: CITY: Abbott Park
: STATE: IL
: COUNTRY: USA
: ZIP: 60064-3500
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/944,483
: FILING DATE:
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Becker, Cheryl L.
: REGISTRATION NUMBER: 35,441
: REFERENCE/DOCKET NUMBER: 6183.US.01
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 847/935-1729
: TELEFAX: 847/938-2623
: TELEX:
: INFORMATION FOR SEQ ID NO: 35:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 224 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: No. 6232456e
: US-08-944-483-35

Alignment Scores:
Pred. No.: 3 1e-30 length: 224
Score: 344.50 Matches: 76
Percent Similarity: 54.25% Conservative: 39
Best Local Similarity: 35.85% Mismatches: 94
Query Match: 22.08% Indels: 3
DB: 3 Gaps: 2

US-10-037-270-482 (1-866) x US-08-944-483-35 (1-224)
Oy 108 AAGAGACCCCTGGCTCCCTATTGGTGTACTGACGTCACCTCAACCCCTGTGGGC 167
Db 8 Glucitanserleuprototyrglnvalserleuansercylserhisphetcysgly 27
Oy 168 GTCTCATCAACCCAGCAGCTGGGTGGTGGCCCGACGTCACCTGCTATTACCAATCTGAAA 227
Db 28 serleuileserglugintyprvalvalseralalalhiserytyrlysthrargliegin 47
Oy 228 GTGATGCTGGGAATTTCAAGAGAGAGTCAGAGAGCGTATCTGAACAGACATTTAAACCC 287
Db 48 Valargleuglygluhisalanillysvalleuglyglucylglninduglnhetleenalala 67

```

[illegible]

```

1      RESULT 9
2      US-08-956-267A-2
3      Sequence 2, Application US/08956267A
4      Patent No. 5945328
5      GENERAL INFORMATION:
6      APPLICANT: WOLDIKE, Helle Fabricius
7      APPLICANT: KJELDSEN, Thomas Borglum
8      TITLE OF INVENTION: A Process For Producing Trypsin
9      TITLE OF INVENTION: (Trypsinogen)
10     NUMBER OF SEQUENCES: 6
11     CORRESPONDENCE ADDRESSES:
12     ADDRESSEE: No. 59453280k of No. 59453280h America, Inc
13     STREET: 405 Lexington Avenue
14     CITY: New York
15     STATE: NY
16     COUNTRY: USA
17     ZIP: 10174
18     COMPUTER READABLE FORM:
19     MEDIUM TYPE: Diskette
20     COMPUTER: IBM Compatible
21     OPERATING SYSTEM: DOS
22     SOFTWARE: FastSeq for Windows Version 2.0
23     CURRENT APPLICATION DATA:
24     APPLICATION NUMBER: US/08/956,267A
25     FILING DATE: 22-Oct-1997
26     CLASSIFICATION: 435
27     ATTORNEY/AGENT INFORMATION:
28     NAME: Rozek, Carol
29     REGISTRATION NUMBER: 36,993
30     REFERENCE/DOCKET NUMBER: 4500.204-US
31     TELECOMMUNICATION INFORMATION:
32     TELEPHONE: 212-867-0123
33     TELEFAX: 212-878-9655
34     INFORMATION FOR SEQ ID NO: 2:
35     SEQUENCE CHARACTERISTICS:
36     LENGTH: 247 amino acids
37     TYPE: amino acid
38     TOPOLOGY: linear
39     MOLECULE TYPE: protein
40     US-08-956-267A-2

```

Alignment Scores:

Pred. No.:	3,376-29	Length:	247
Score:	335.50	Matches:	78
Percent Similarity:	52.89%	Conservative:	41
Best Local Similarity:	34.67%	Mismatches:	97
Query Match:	21.51%	Indels:	9
DB:	2	Gaps:	2

US-10-037-270-482 (1-866) x US-08-956-267A-2 (1-247)

```

QY 69 CTCGCTGGGACATTTTCTTCTGCTGACTGCTGTTCAAGAAAGAACCTGCTCCCTAT 128
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25 ILevalGlyGlyThrCysAlaAlaAsnSerIle-----ProTyr 38
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 129 TTGGTGTACTCAAGCTCTCACTTCAACCCCTGTGGGCTCTCTCAATCAACCCAGTGG 188
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 39 GlnValSerLeuAsnSerGlySerHisPheCysGlyGlySerLeuIleAsnSerGlnTyr 58
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 189 GTGCTGCCCCCAGCTCACTGCTATTTACCAATCTGAAAGTGAATGCTGGAAATTTCAAG 248
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 59 ValValSerAlaAlaHisCysTyrTyrSerArgIleGlnValArgLeuGlyLysAsn 78
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 249 AGCAGAGTCAGACGCTGACTGACAGACAAATTACCCCATTCAGATGCTCGCTACTGG 308
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 79 IleAspValLeuGlyLysGlnGlnPheIleAsnAlaAlaIleThrHisPro 98
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 309 AACTACAGTCATAGCCGCCCAAGATGACTGATGCTCAATCAAGCTGCTAGCTGCC 368
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 99 AsnPheAsnGlyAsnThrLeuAspAsnAspIleMetLeuIleLysLeuSerSerProAla 118
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 369 ATGCTCAATCCCAAGCTCCAGCCCTTCCCTCGCCCAACCAATGACAGCCAGCACT 428
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 119 ThrLeuAsnSerArgValAlaThrValSerLeuProArgSerCysAlaAlaAlaGlyThr 138
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 429 GTCTGTCTACTCTCAGGTTTGAATGAGCCAGCAAAAAACAGTGGCCGACACCTGACTTG 488
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 139 GluCysLeuIleSerGlyTyrGlyAsnThrLysSerSerCysSerTyrProSerLeu 158
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 489 CGGCAGAACCTGAGAGCCGCCGATGCTGATCGAAGATCCAAAAACAGAACAGAA 548
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 159 LeuGlnCysLeuLysAlaProValIleuSerAspSerSerCysLysSerSerTyrProGly 178
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 549 AAAAGCCACAGAAATCTTATGTTGAATTTGTGAAGATTTTCAAGCCCAATTTTGGG 608
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 179 GlnIleThrGlyAsnMetIleCysValGlyPheLeuGlnGlyLysAspSerCysGln 198
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 609 GAGGTGCGCTGCTCTACTGCTCATCTGCAAGACAGCAAGCTCCAGGGAATC-----GAG 659
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 199 GlyAspSerGlyGlyProValValCysAsnGlyGlnLeuGlnGlyIleValIleSerTyrGly 218
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 660 GTGGGCACTTCAATGGAGGGGAGCTGCGCATTCACCAATTTTCAAAATATGATCC 719
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 219 TyrGlyCysAlaGlnLysAsnLysProGlyValIleThrLysValCysAsnTyrValAsn 238
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 720 TGGATTGAGAACT 734
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 239 TrpIleGlnGlnThr 243
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
US-08-944-483-34
; Sequence 34, Application US/08944483
; Patent No. 6232456
; GENERAL INFORMATION:
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: KASS, MICHAEL R.
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STEWART, KENT D.
; APPLICANT: STROUPE, STEVEN D.
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS

```

```

; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
; TITLE OF INVENTION: OF THE PROSTATE
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,483
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6183.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 224 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: NO. 62324566
; US-08-944-483-34

Alignment Scores:
Pred. No.: 4,176-29 Length: 224
Score: 334.50 Matches: 75
Percent Similarity: 52.83% Conservative: 37
Best Local Similarity: 35.38% Mismatches: 97
Query Match: 21.44% Indels: 3
DB: 3 Gaps: 1

US-10-037-270-482 (1-866) x US-08-944-483-34 (1-224)

QY 108 AAAGAAGACCCGCTCCCTATTTGCTGATCTCAATGCTCAACCCCTGTGGGC 167
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8 GluGlnAsnSerValProTyrGlnValSerLeuAsnSerGlyTyrHisPheCysGlyGly 27
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 168 GTCCCTCATCAAAACCCAGCTGGGTGCGCCCAAGCTCACTGCTATTTTCAAAATGAAA 227
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 28 SerLeuIleSerGlnGlnTyrValValSerAlaGlyHisCysTyrTyrSerArgIleGln 47
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 228 GTGATGCTGGAAATTTCAAGAGCAGAGTCAGAGCGTACTGACAGCAAGCAATTAACCCC 287
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 48 ValArgLeuGlyGlyHisAsnIleGluValLeuGlnGlyAsnGlnGlnPheIleAsnAla 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 288 ATTCAAGATGCTCCGCTACTGAGAACTCACTCATAGGCCCCCAAGATACCTCAAGTCC 347
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68 AlaValIleIleArgHisProLysTyrAsnSerArgThrLeuAspAsnAspIleLeu 87
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 348 ATCAAGCTGAGTAAAGCTCCATGCTCAATCCCAAGTCCAGCCCTTCCCTCGCCACC 407
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 88 IleLysLeuSerSerProAlaValIleAsnSerArgValSerAlaIleSerLeuProThr 107
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 408 ACCAATGCAAGCCAGCACTGTCTGTACTCTCAAGTTTGAAGTGAAGCAAGCAAAAG 467
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108 AlaProProAlaAlaGlyThrGlnSerLeuIleSerGlyTyrGlyAsnThrLeuSerSer 127
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


APPLICATION NUMBER: US/08/628,198
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/467,155
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Clark, Paul T.
 REGISTRATION NUMBER: 30,162
 REFERENCE/DOCKET NUMBER: 00398/100002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617/542-5070
 TELEFAX: 617/542-8906
 TELEX: 200154
 INFORMATION FOR SEQ. ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 281 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-628-198-7

Alignment Scores:
 Pred. No.: 2,29e-27 Length: 281
 Score: 319.50 Matches: 81
 Percent Similarity: 49.79% Conservative: 40
 Best Local Similarity: 33.33% Mismatches: 91
 Query Match: 20.48% Indels: 31
 Gaps: 6

US-10-037-270-482 (1-866) x US-08-628-198-7 (1-281)

QY 108 AAAGAAGACCTGCTCCCTATTGTTGGTACCTCACTCAATCCCTGTGGGC 167
 DB 38 ArgGluSerSerValProTyrGlnValSerLeuAsnAlaGlyTyrHisPheCysGlyGly 57
 QY 168 GTCCCTCATCAAAACCCAGCTGGGCTGGCCGCCCTCACTCACTTATTCAAATCGA 227
 DB 58 SerLeuIleAsnAspIleTyrValValSerAlaAlaHisCysTyrTyrArgIleGln 77
 QY 228 GTGATCTGTGGGA-----AATTTCAAGACGAGA-----GTCAGACAGCGT 266
 DB 78 ValArgLeuGlyGlnHisAsnMetMetThrArgTyrAlaArgIleAsnValLeuGluGly 97
 QY 267 ACTGAACAGACAATTAACCCCATTCAGATGTCGCTACTGAGCACTACACTATTAACGCC 326
 DB 98 AsnGluGlnPheValAspSerAlaValIleIleArgHisProAsnTyrAsnSerTrpThr 117
 QY 327 CCACAGAGATGACTCATGCTCAAGCTGGCTAAGCTTGCCTGCTC----- 374
 DB 118 LeuAspAsnAspIleMetLeuIleValSerProAlaThrLeuMetMetThrArg 137
 QY 375 -----AATCCCAAGTCAGCCCTTCCCTCGCCACCAATGTCAGCCAGGC 425
 DB 138 TyrAlaArgAsnAlaArgValAlaSerValProLeuProSerSerCysAlaProAlaGly 157
 QY 426 ACTGTCTGTACTCTCAGGTTTGACTGAGCCCAAGAAAAGAGGCGCCACCTGAC 485
 DB 158 ThrGlnCysLeuIleSerGlyTyrGlyAsnThrLeuSerAsnGlyValAsnAsnProAsp 177
 QY 486 TTCCGGCAGAACCTGAGAGCCCTCGATGTCGTGAT-----CGA 524
 DB 178 LeuLeuGlnCysValAspAlaProValProGlnAlaMetMetThrArgTyrAlaArg 197
 QY 525 GAATGCCAAATAAAGACAGAAAGAAAAGCCACAGAAATTCCTATGTGTAATTTGTG 584
 DB 198 AspCysGlnAlaSerTyrProGlyAspIleThrAsnAsnMetIleCysValGlyPheLeu 217
 QY 585 AAAGTATTCAGCCGAATTTTGGGAGAGTGCCGTGCTACTGCTACCTGCAAAAGCAAG 644
 DB 218 GluGlyGlyIysAspSerCysGlnGlyAspSerGlyGlyProValValCysAsnGlyGlu 237

QY 645 CTCACAGGAATC---GAGGTGGGCACTTCATG----- 674
 DB 238 LeuGlnGlyIleValSerTrpGlyTyrMetMetThrArgTyrAlaArgGlyCysAlaGln 257
 QY 675 GGAGGGAGCTGGCATCTACCAATGTTTCAATATGATTCCTGGATTGAGAACT 734
 DB 258 ProAspAlaProGlyValTyrThrIysValCysAsnTyrValAspTrpIleGlnAsnThr 277
 QY 735 GCTAAGGAC 743
 DB 278 IleAlaAsp 280

RESULT 13

US-09-201-038-7
 Sequence 7, Application US/09201038
 Patent No. 6153387
 GENERAL INFORMATION:
 APPLICANT: Band, Viola
 TITLE OF INVENTION: NES-1 POLYPEPTIDES, DNA, AND RELATED
 TITLE OF INVENTION: MOLECULES AND METHODS
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fien & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/201,038
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/628,198
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Clark, Paul T.
 REGISTRATION NUMBER: 30,162
 REFERENCE/DOCKET NUMBER: 00398/100002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617/542-5070
 TELEFAX: 617/542-8906
 TELEX: 200154
 INFORMATION FOR SEQ. ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 281 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-201-038-7

Alignment Scores:
 Pred. No.: 2,29e-27 Length: 281
 Score: 319.50 Matches: 81
 Percent Similarity: 49.79% Conservative: 40
 Best Local Similarity: 33.33% Mismatches: 91
 Query Match: 20.48% Indels: 31
 Gaps: 6

US-10-037-270-482 (1-866) x US-09-201-038-7 (1-281)

QY 108 AAAGAAGACCTGCTCCCTATTGTTGGTACCTCACTCAATCCCTGTGGGC 167
 DB 38 ArgGluSerSerValProTyrGlnValSerLeuAsnAlaGlyTyrHisPheCysGlyGly 57
 QY 168 GTCCCTCATCAAAACCCAGCTGGGCTGGCCGCCCTCACTCACTTATTCAAATCGA 227

```

Db      58 SerLeuIIeAsnAspIInTrpValIaSerAlaAlaHisCysTrpIlySerTyArgIleGln  77
QY      228  GNGATGCTGGGA-----AATTCCAAGAGAGAG-----GTCAGAGACGGT  266
Db      78  ValAArgLeuGIgLIuHisAsnMetMetThrArgTyAlaArgIleAsnValLeuGIuGly  97
QY      267  ACTGAACAAGACAAATTAAACCCATTCAATGCATGCCTCCGCTACTGAGAACTACATCAATGACGCC  326
Db      98  AaNGIuGInPheValaAspSerAlaValSleIleIleArgHisProAsnTyIyrAsnSerTrpThr  117
QY      327  CCACAGGATGACCTCATCTCATCAAGCTGGCTAGCCCTGCATGCTC-----  374
Db      118  LeuAspAsnAspIleMetLeuIleIySleuAlaSerProValThrIleuMetMetThrArg  137
QY      375  -----AATCCCAAGATCCAGCCCCCTTCCCTCGCCACCAACATGTCCAGCCAGCC  423
Db      138  TyAlaArgAsnAlaArgValaIaSerValProIeuProSerSerCysAlaProIaGly  157
QY      426  ACTGTCTGTCTACTCTCAGATTTTGACTGGAGCCAGAAACAGTGCCGACACCTTGAC  485
Db      158  ThrGInCysIeuIleSerGIlyTrpGIyAsnThrIleuSerAsnGlyValaAsnAsnProAsp  177
QY      486  TTGCGGACAGAACCTCAGAGGCCCCCGGTGATGTCTGAT-----CGA  524
Db      178  LeuIeuGInCysValaAspAlaProValaIeuProGInAlaIeMetMetThrArgTyAlaArg  197
QY      525  GAATGCCCAAAAACAGAACAGAAAAGCCACAGAAATCTTTATGTGTGAATTGTG  584
Db      198  AspCysGIuAlaSerTyIyrProGIyAspIleThrAsnAsnMetIleCysValGIyPheIeu  217
QY      585  AAGATTTACGCCGAATTTTGGCGAGGAGGCCGCTGCTACTGTCATCTGCCAAAGACAG  644
Db      218  GIuGIyGIyIyAspSerCysGInGIyAspSerGIyIyProValaIyCysAaNGIyIu  237
QY      645  CTCGAGGGAATC---GAGGTGGGGCAGCTTCATG-----  674
Db      228  LeuGIuGIyIleValaSerTrpGIyIyTrpMetMetThrArgTyAlaArgIyCysAlaGIn  257
QY      675  GGAGGGAGACGTGGCATCTACACCAATGTTTACAAATATGTATCTGGATTGGAACACT  734
Db      258  ProAspAlaProGIyValIyTrpIyIyAsnTyValaAsnTyIyrIleGIuAsnThr  277
QY      735  GCTAAGGAC  743
Db      278  IleAlaAsp  280

RESULT 14
PCT-US96-07343-7
: Sequence 7, Application PC/TUS9607343
: GENERAL INFORMATION:
: APPLICANT: New England Medical Center Hospitals, Inc.
: TITLE OF INVENTION: NES-1 POLYPEPTIDES, DNA, AND RELATED
: TITLE OF INVENTION: MOLECULES AND METHODS
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02110-2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US96/07343
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/467,155
: FILING DATE: 06-JUN-1995
:

```

```

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00398/100001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 281 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-07343-7

Alignment Scores:
Pred. No.: 2,29e-27 Length: 281
Score: 319.50 Matches: 81
Percent Similarity: 49.79% Conservative: 40
Best Local Similarity: 33.33% Mismatches: 91
Query Match: 20.48% Indels: 31
DB: 5 Gaps: 6

US-10-037-270-482 (1-866) x PCT-US96-07343-7 (1-281)

Oy 108 AAGAAGAACCCCTGCTCCCTATTGGTGATCCTCAAGTCTCAACCCCTGTGTGGC 167
Db 38 ArgGlnSerSerValProTyrGlnValSerLeuAsnAlaGlyTyrHisPheCysGly 57
Oy 168 GTCCATCAAAACCCAGCTGGGTGTGGCCGCCAGCTCACTGTAATTTACCAATGTGAA 227
Db 58 SerLeuIleAsnAspIleTyrValValSerAlaAlaHisCysTyrTyrArgIleGln 77
Oy 228 GTGATGCTGGG-----AATTCAAGACGAG-----GTACGACAGCGT 266
Db 78 ValArgLeuGlyValHisAsnMetThrArgTyrAlaArgIleAsnValLeuGlnGly 97
Oy 267 ACTGAACAGACATTAAACCCCATTCAGATCGCGCTACCTGAACTACAGTCATAGCGCC 326
Db 98 AsnGlnGlnPheValAspSerAlaValIleIleArgHisProAsnTyrAsnSerThr 117
Oy 327 CCACAGATGACCTCATGCTCATCAAGCTGTGCTAGCCCTGCCATGCTC----- 374
Db 118 LeuAspAsnAspIleMetLeuIleTyrLeuAlaSerProValThrLeuMetThrArg 127
Oy 375 -----AATCCAAATCCAGCCCTTCCTCCCTCGCACACCAATGTCAGGCCAGGC 425
Db 138 TyrAlaArgAsnAlaArgValAlaSerValProLeuProSerSerCysAlaProAlaGly 157
Oy 426 ACTGTGTGTCTACTCTCAGATTGTGACAGCGCAAGAAAACAAGTGGCCGACACCTTGAC 485
Db 158 ThrGlnCysLeuIleSerIleTyrPheValAsnThrLeuSerAsnGlyValAsnAspAsp 177
Oy 486 TTGCGACAGAACCTGTAGGCCCCCTGTAGTCTGAT-----CGA 524
Db 178 LeuLeuGlnCysValAlaAspAlaProValLeuProGlnAlaMetThrArgTyrAlaArg 197
Oy 525 GAATGCCAAAAACAGAACAGAAAACCAACAGAAATTCCTATGTGTGAATTTGTG 584
Db 198 AspCysGlnAlaSerTyrProIleAspIleThrAsnAsnMetIleCysValGlyPheLeu 217
Oy 585 AAAGTATTACGCCGAATTTTGGGAGAGTGGCCCTGTCTACTGTCACTGTCCAAAGACAG 644
Db 218 GlnGlyGlyLysAspSerCysGlnGlyAspSerGlyGlyProValValCysAsnGly 237
Oy 645 CTCACAGGATC---GAGGTGGGACCTTCACTG----- 674
Db 238 LeuGlnGlyIleValSerTyrPheIleThrArgTyrAlaArgGlyCysAlaGln 257
Oy 675 GAGGGGAGCGTGGGATCTACCAACAAATGTATTCCTGATTTGAGAACACT 724

```

DB 258 ProAspAlaProGlyValTyrThrLysValCysAsnTyrValAspTrpIleGlnAsnThr 277
QY 735 GCTAAGAC 743
DB 278 IleAlaAsp 280

RESULT 15
US-08-278-091-9
Sequence 9, Application US/08278091
Patent No. 5506139
GENERAL INFORMATION:
APPLICANT: LOOSMORE, Sheena M
APPLICANT: YANG, Yan-Ping
APPLICANT: CHONG, Pele
APPLICANT: COWEN, Raymond P.
APPLICANT: KLEIN, Michel H.
TITLE OF INVENTION: Analog of Haemophilus Hln47 Protein with
TITLE OF INVENTION: Reduced Protease Activity
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/278, 091
FILING DATE: 21-JUL-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 223 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-278-091-9

Alignment Scores:
Pred. No.: 1,27e-26 Length: 223
Score: 312.50 Matches: 71
Percent Similarity: 52.66% Conservative: 38
Best Local Similarity: 34.30% Mismatches: 95
Query Match: 20.03% Indels: 3
DB: 1 Gaps: 1

US-10-037-270-482 (1-866) x US-08-278-091-9 (1-223)

QY 123 CCTATTGTGTTACTCAAGTCTCACTTCAACCCGCTGTGGCGCTCTCATCAACC 182
DB 13 PctyTclnAlaSerLeuAsnSerGlyTyrHisPheCysGlySerLeuIleAsnSer 32
QY 183 AGCTGGTGTGCTGCCAGCTCACTGCTATTACCAATCTGAAGTATGCTGGAAAT 242
DB 33 GlnTrpValValSerAlaAlaHisCysTyrLysSerGlyIleGlnValArgLeuGlyGln 52
QY 243 TTCAGAGCAGAGTCAAGACGGTACTGAACAGACCAATTACCCCATTCAGATCGTCCG 302
DB 53 AspAsnIleAsnValValGlnGlyAsnGlnGlnPheIleSerAlaSerLysSerIleVal 72

QY 303 TACTGGAAGTACATGATAGCCGCCACAGATGAGTCACTCATGCTGAGCTGCTAG 362
DB 73 HisProSerTyrAsnSerAsnThrLeuAsnAsnAlaIleMetLeuIleLysLeuLysSer 92
QY 363 CCTGCCATGCTCAATCCAAAGTCCAGGCCCTTCCCTGCGCACCAATGTCAGGCCA 422
DB 93 AlaAlaSerLeuAsnSerArgValAlaSerIleSerLeuProThrSerCysAlaSerAla 112
QY 423 GGCACGTGCTGCTCTCTCAGGTTTGGACTGGAGCCAGAAAGTGGCCGACACCCCT 482
DB 113 GlyThrGlnCysLeuIleSerGlyTyrGlyAsnThrLysSerSerGlyThrSerTyrPro 132
QY 483 GACTTCGGCAGAACTTGAGGAGCCCTGATGCTGATGAGAAATGCCAAAAACAGAA 542
DB 133 AspValIleuLysCysLeuLysAlaProIleLeuSerAspSerSerLysSerAlaTyr 152
QY 543 CAAGGAAAAAGCCACAGAAATTCCTTATGTGTGAATTGTGAAAGTATTACGCCAATT 602
DB 153 ProGlyGlnIleThrSerAsnMetPheCysAlaGlyTyrLeuGlnGlyLysAspSer 172
QY 603 TTGGGAGGTGGCGCTTCTACTGTCATCTGCACAAAGCTCCAGGAATC----- 656
DB 173 CysGlnGlyAspSerSerGlyProValValCysSerGlyLysLeuGlnGlyIleValSer 192
QY 657 ---GAGGTGGGACACTTCATGGAGGAGGAGCTGCGCATCTACACCAATGTTTACAATAT 713
DB 193 TrpGlySerGlyCysAlaGlnLysAsnLysProGlyValTyrThrLysValCysAsnTyr 212
QY 714 GATTCCTGATGAGAACACT 734
DB 213 ValSerTrpIleLysGlnThr 219

Search completed: November 8, 2003, 02:13:40
Job time : 27 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 8, 2003, 02:11:47 ; Search time 181.5 Seconds
(without alignments)
1638.952 Million cell updates/sec

Title: US-10-037-270-482
Perfect score: 1560
Sequence: 1 ggcactactccctgagcta.....tgggaatgagcatcactagt 866

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 644079 segs, 171749292 residues

Total number of hits satisfying chosen parameters: 1288158

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=x1h
-Q=/cgn2_1/USFTO.epool/US10037270/runat.07112003.140520.8870/app_query.fasta_1.1031
-DB=Published Applications AA -QFMT=fastan -SUFFIX=rapp -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.csi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10037270 @CGN_1.1.50 @runat.07112003.140520.8870
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:*

1:	/cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/2/pubppa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/2/pubppa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/2/pubppa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/2/pubppa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pep.*
16:	/cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep.*
17:	/cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep.*
18:	/cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	1255	80.4	235	11	US-09-842-758-12	Sequence 12, App1
2	1255	80.4	235	11	US-09-842-758-14	Sequence 14, App1
3	1250	80.1	235	10	US-09-910-151-2	Sequence 2, App1
4	1247	79.9	235	12	US-10-246-006-2	Sequence 102, App1
5	1222	78.3	265	9	US-09-888-615-102	Sequence 107, App
6	1217	78.0	231	11	US-09-842-758-107	Sequence 53, App1
7	377.5	24.2	248	11	US-09-842-758-54	Sequence 53, App1
8	375.5	24.1	248	11	US-09-842-758-55	Sequence 154, App
9	351.5	22.5	246	11	US-09-842-758-55	Sequence 55, App1
10	334.5	21.4	247	9	US-09-823-779-154	Sequence 50, App1
11	328	21.0	215	11	US-09-898-837A-30	Sequence 7, App1
12	319.5	20.5	281	14	US-10-021-368-7	Sequence 14, App1
13	312.5	20.0	223	10	US-09-910-071-14	Sequence 39, App1
14	311	19.9	243	12	US-10-239-663-39	Sequence 248, App
15	305	19.6	241	12	US-10-137-870-248	Sequence 248, App
16	305	19.6	241	12	US-10-140-018-248	Sequence 248, App
17	305	19.6	241	12	US-10-140-021-248	Sequence 248, App
18	305	19.6	241	12	US-10-140-274-248	Sequence 248, App
19	305	19.6	241	12	US-10-140-471-248	Sequence 248, App
20	305	19.6	241	12	US-10-140-807-248	Sequence 248, App
21	305	19.6	241	12	US-10-140-922-248	Sequence 248, App
22	305	19.6	241	12	US-10-140-922-248	Sequence 248, App
23	305	19.6	241	12	US-10-140-926-248	Sequence 248, App
24	305	19.6	241	12	US-10-141-698-248	Sequence 248, App
25	305	19.6	241	12	US-10-141-702-248	Sequence 248, App
26	305	19.6	241	12	US-10-141-704-248	Sequence 248, App
27	305	19.6	241	12	US-10-141-704-248	Sequence 248, App
28	305	19.6	241	12	US-10-142-432-248	Sequence 248, App
29	305	19.6	241	12	US-10-142-767-248	Sequence 248, App
30	305	19.6	241	12	US-10-142-767-248	Sequence 248, App
31	305	19.6	241	12	US-10-143-033-248	Sequence 248, App
32	305	19.6	241	12	US-10-144-994-248	Sequence 248, App
33	305	19.6	241	12	US-10-145-628-248	Sequence 248, App
34	305	19.6	241	12	US-10-145-631-248	Sequence 248, App
35	305	19.6	241	12	US-10-145-633-248	Sequence 248, App
36	305	19.6	241	12	US-10-145-748-248	Sequence 248, App
37	305	19.6	241	12	US-10-145-748-248	Sequence 248, App
38	305	19.6	241	12	US-10-145-823-248	Sequence 248, App
39	305	19.6	241	12	US-10-145-826-248	Sequence 248, App
40	305	19.6	241	12	US-10-145-870-248	Sequence 248, App
41	305	19.6	241	12	US-10-145-876-248	Sequence 248, App
42	305	19.6	241	12	US-10-145-959-248	Sequence 248, App
43	305	19.6	241	12	US-10-146-724-248	Sequence 248, App
44	305	19.6	241	12	US-10-146-725-248	Sequence 248, App
45	305	19.6	241	12	US-10-147-435-248	Sequence 248, App

ALIGNMENTS

RESULT 1
US-09-842-758-12
; Sequence 12, Application US/09842758
; Publication NO. US20030083244A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Majumder, Kumud
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patlurajan, Meera
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gangoli, Reha A.
; APPLICANT: Smithson, Glenda
; APPLICANT: Rastelli, Luca
; APPLICANT: MacDougall, John R.
; APPLICANT: Taupier, Raymond J.
; APPLICANT: Grose, William M.

APPLICANT: Edward, Szekeres S
APPLICANT: Alsobrook II, John P
TITLE OF INVENTION: No. US20030083244A1el Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 15966-783
CURRENT FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: US/09/842,758
PRIOR FILING DATE: 2000-04-26
PRIOR APPLICATION NUMBER: 60/200,158
PRIOR FILING DATE: 2000-04-26
PRIOR APPLICATION NUMBER: 60/200,613
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,780
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/201,006
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,007
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,236
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,238
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,186
PRIOR FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: 60/201,474
PRIOR FILING DATE: 2000-05-03
PRIOR APPLICATION NUMBER: 60/201,508
PRIOR FILING DATE: 2000-05-03
PRIOR APPLICATION NUMBER: 60/220,591
PRIOR FILING DATE: 2000-07-25
PRIOR APPLICATION NUMBER: 60/232,678
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 60/263,217
PRIOR FILING DATE: 2001-01-22
PRIOR APPLICATION NUMBER: 60/265,160
PRIOR FILING DATE: 2001-01-30
NUMBER OF SEQ ID NOS: 113
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 235
TYPE: PRT
ORGANISM: Homo sapiens
US-09-842-758-12

Alignment Scores:
Pred. No.: 2e-116 Length: 235
Score: 1255.00 Matches: 235
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 80.45% Indels: 0
DB: 11 Gaps: 0

US-10-037-270-482 (1-866) x US-09-842-758-12 (1-235)

QY 42 ATGAATATGCTCTTATTTGGGTGCTCGCGGACATTTTCTTGTGCTACTACT 101
DB 1 MetLeuIeYValPheTYrLeuGlyValLeuAlaGlyThrPhePheAlaAepSerSer 20
QY 102 GTTCAGAAAGAACCCCTGCTCCCTATTTGGTGAAGTCAAGTCACTTCAACCCCTGT 161
DB 21 ValGlnIysGlnuSprProAlaProTYrLeuValTYrLeuYsSerHIsPheAsnProCys 40
QY 162 GTGGGCGTCTCTCATCAACCCAGCTGGGTGCTGGCCCGCCAGCTCACTGCTATTACCAAT 221
DB 41 ValGlyValLeuIleYsProSerTrpValLeuAlaProAlaHIsCysTYrLeuProAsn 60
QY 222 CTGAAGTAGTGTGGGAAATTTCAAGACAGACTGAGACGCTACTGAACAGCAAT 281
DB 61 LeuIysValMetLeuGlnYAsnPheYsSerArgValArgAspGlyThrGlnGlnThrIle 80
QY 282 AACCCCTTGAATCGTCCGCTACTGGAATCAAGTATAGCGCCCAAGATGATGACTC 341
DB 81 AsnProIleGlnIleValArgTYrTrpAsnTYrSerHIsSerAlaProGlnAspAspLeu 100
QY 342 ATGCTCATCAAGCTGGCTAAGCTGCATGCTCAATCCCAAGTCAGCCCTTCCCTC 401

DB 101 MetLeuIleYsValLeuAlaYsProAlaMetLeuAsnProCysValGlnProLeuProLeu 120
QY 402 GCCACACCAACCAATTCACGCGCCAGGCACTGCTGTCTACTCTCAGATTGACCTGAGCCAA 461
DB 121 AlaThrThrAsnValArgProGlyThrValCysLeuIleSerTrpLeuAsnProSerGln 140
QY 462 GAAACAGTGGCGGACCACTGACTGCGGACAGAACCTGAGGCGCCCGCTGATGTGAT 521
DB 141 GlnAsnSerTrpYrHIsProAspLeuArgGlnAsnLeuGlnAlaProValMetSerAsp 160
QY 522 CGAAGATGCCAAAAACAGAACAGGAAAAAGCCACAGCAATTCCTTATGTGAAATTT 581
DB 161 ArgGlnCysGlnIysThrGlnGlnIysSerHIsArgAsnSerLeuCysValYsPhe 180
QY 582 GTGAAGTATTCAGCCGAATTTTGGGAGAGTGGCCGCTGCTCATCTGCAACCAAG 641
DB 181 ValIysValPheSerArgIlePheGlyGlnValAlaValAlaThrValIleCysYsAsp 200
QY 642 AACCTCAGGGAATTCAGGTGGGCACTTCATGAGAGGGAAGCTCGGCAATCTACCAAT 701
DB 201 LysLeuGlnIysIleGlnValAlaGlyHIsPheMetGlyGlyAspValGlyIleTYrThrAsn 220
QY 702 GTTACAAATATGTAATCTGGATTGAGAACTGCTTAAGGACAAAG 746
DB 221 ValTYrIysTYrValSerTrpIleGlnuAsnThrAlaYsAspYs 235

RESULT 2
US-09-842-758-14
Sequence 14, Application US/09842758
Publication No. US20030083244A1
GENERAL INFORMATION:
APPLICANT: Vernet, Corine A. M.
APPLICANT: Bernades, Elma R.
APPLICANT: Gerlach, Valerie
APPLICANT: Shinkets, Richard A
APPLICANT: Malynkar, Uriel M
APPLICANT: Boldog, Ferenc L
APPLICANT: Zerusen, Bryan D
APPLICANT: Spytek, Kimberly A
APPLICANT: Majumder, Komud
APPLICANT: Tchener, Velizar T
APPLICANT: Padigaru, Muralidhara
APPLICANT: Patirajan, Meera
APPLICANT: Burgess, Catherine E
APPLICANT: Gangoli, Beba A
APPLICANT: Smithson, Glenda
APPLICANT: Rastelli, Luca
APPLICANT: MacDougall, John R
APPLICANT: Taupier, Raymond J
APPLICANT: Grose, William M
APPLICANT: Edward, Szekeres S
APPLICANT: Alsobrook II, John P
TITLE OF INVENTION: No. US20030083244A1el Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 15966-783
CURRENT APPLICATION NUMBER: US/09/842,758
CURRENT FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: 60/200,158
PRIOR FILING DATE: 2000-04-26
PRIOR APPLICATION NUMBER: 60/200,613
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,780
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/201,006
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,007
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,236
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,238
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,186
PRIOR FILING DATE: 2000-05-02

PRIOR APPLICATION NUMBER: 60/201,474
PRIOR FILING DATE: 2000-05-03
PRIOR APPLICATION NUMBER: 60/201,508
PRIOR FILING DATE: 2000-05-03
PRIOR APPLICATION NUMBER: 60/220,591
PRIOR FILING DATE: 2000-07-25
PRIOR APPLICATION NUMBER: 60/232,678
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 60/263,217
PRIOR FILING DATE: 2001-01-22
PRIOR APPLICATION NUMBER: 60/265,160
PRIOR FILING DATE: 2001-01-30
NUMBER OF SEQ ID NOS: 113
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 14
LENGTH: 235
TYPE: PRT
ORGANISM: Homo sapiens
US-09-842-758-14

Alignment Scores:

Pred. No.:	2e-116	Length:	235
Score:	1255.00	Matches:	235
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	80.45%	Indels:	0
DB:	11	Gaps:	0

US-10-037-270-482 (1-866) x US-09-842-758-14 (1-235)

42 ATGAATATGCTCTTATTTGGGTGCTCGCTGGAGACATTTTCTTGGTCACTATCT 101
1 MetLysTyrValPheTyrLeuGlyValLeuAlaGlyThrPhePheAlaSerSer 20
102 GTTCAGAAAGAGACCTGCTCTCTATTTGGTGTACCTCAAGTCACTTCAACCCCTGT 161
21 ValGlnLysGlnAspProAlaProTyrLeuValTyrLeuLysSerHisPheAsnProCys 40
162 GTGGGGCTCTCTATCAAAACCCAGCTGGGTGCTGGCCCACTACGCTATTATCCAAAT 221
41 ValGlyValLeuLysProSerTrpValLeuAlaProAlaHisCysTyrLeuProAsn 60
222 CTGAAGTGTGCTGGAAATTTCAAGACAGAGTCAAGACCGTACTGAACAGACAATT 281
61 LeuLysValMetLeuGlyAsnPheLysSerArgValArgAspLysTrpGlnInthrIle 80
282 AACCCCATTCAGATGCTCGCTACTGGAATCAAGTCAATGACGCCCAAGATGACCTTC 341
81 AsnProIleGlnIleValArgTyrTrpAsnTyrSerHisSerAlaProGlnAspLeu 100
342 ATGCTCATCAAGCTGGCTAAGCTGCTGATGCTCAATCCCAAGTCCAGCCCTTCCCTTC 401
101 MetLeuLysValLeuAlaLysProAlaMetLeuAsnProLysValGlnProLeuProLeu 120
402 GCCACCAACCAATGTCAGGCGGACGCTGTCTGCTCACTGCTGAGTTGGACTGGAGCAA 461
121 AlaThrThrAsnValArgProGlyThrValCysLeuLeuSerGlyLeuAspTrpSerGln 140
462 GAAACAGTGGCCGACACCTGACTTGGCGGACAGACCTGGAGGCCGCCGCTGATGCTGAT 521
141 GlnAsnSerGlyArgHisProAspLeuArgGlnAsnLeuGlnAlaProValMetSerAsp 160
522 CGAGAAATGCCAAAAACAGAACAGAAAAAGCCACAGAAATCTTATGTGTAATTT 581
161 ArgGlnCysGlnLysThrGlnGlnGlyLysSerHisArgAsnSerLeuCysValLysPhe 180
582 GTAAGATTTTACCGCAATTTTGGGAGGTGGCGGTTGCTACTGCTACATGCAAGAC 641
181 ValLysValPheSerArgLysPheGlyValAlaValAlaThrValIleCysLysAsp 200
642 AACCTCAGAGAAATCAGATGGGCACTTCATGGAGGAGGAGCTCGCATCTTACACCAAT 701
201 LysLeuGlnGlnGlyLysGlyValGlnLysPheMetCysGlyAspValGlyIleTyrThrAsn 220

702 GTTACAAATATGATCTCTGATTTGAGAACACGCTAAGACAAAG 746
221 ValTyrLysTyrValSerTrpIleGlnAsnThrAlaLysAspLys 235

RESULT 3

US-09-910-151-2
Sequence 2, Application US/09910151
Patent No. US20020137181A1
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel
TITLE OF INVENTION: 14087, A NOVEL SERINE PROTEASE MOLECULE
TITLE OF INVENTION: AND USES THEREFOR
FILE REFERENCE: 38155-20021.00
CURRENT APPLICATION NUMBER: US/09/910,151
CURRENT FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: US 60/219,002
PRIOR FILING DATE: 2000-07-17
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 235
TYPE: PRT
ORGANISM: Homo sapiens
US-09-910-151-2

Alignment Scores:

Pred. No.:	6.31e-116	Length:	235
Score:	1250.00	Matches:	234
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.57%	Mismatches:	0
Query Match:	80.13%	Indels:	0
DB:	10	Gaps:	0

US-10-037-270-482 (1-866) x US-09-910-151-2 (1-235)

42 ATGAATATGCTCTTATTTGGGTGCTCGCTGGAGACATTTTCTTGGTCACTATCT 101
1 MetLysTyrValPheTyrLeuGlyValLeuAlaGlyThrPhePheAlaSerSer 20
102 GTTCAGAAAGAGACCTGCTCTCTATTTGGTGTACCTCAAGTCACTTCAACCCCTGT 161
21 ValGlnLysGlnAspProAlaProTyrLeuValTyrLeuLysSerHisPheAsnProCys 40
162 GTGGGGCTCTCTATCAAAACCCAGCTGGGTGCTGGCCCACTACGCTATTATCCAAAT 221
41 ValGlyValLeuLysProSerTrpValLeuAlaProAlaHisCysTyrLeuProAsn 60
222 CTGAAGTGTGCTGGAAATTTCAAGACAGAGTCAAGACCGTACTGAACAGACAATT 281
61 LeuLysValMetLeuGlyAsnPheLysSerArgValArgAspLysTrpGlnInthrIle 80
282 AACCCCATTCAGATGCTCGCTACTGGAATCAAGTCAATGACGCCCAAGATGACCTTC 341
81 AsnProIleGlnIleValArgTyrTrpAsnTyrSerHisSerAlaProGlnAspLeu 100
342 ATGCTCATCAAGCTGGCTAAGCTGCTGATGCTCAATCCCAAGTCCAGCCCTTCCCTTC 401
101 MetLeuLysValLeuAlaLysProAlaMetLeuAsnProLysValGlnProLeuProLeu 120
402 GCCACCAACCAATGTCAGGCGGACGCTGTCTGCTCACTGCTGAGTTGGACTGGAGCAA 461
121 AlaThrThrAsnValArgProGlyThrValCysLeuLeuSerGlyLeuAspTrpSerGln 140
462 GAAACAGTGGCCGACACCTGACTTGGCGGACAGACCTGGAGGCCGCCGCTGATGCTGAT 521
141 GlnAsnSerGlyArgHisProAspLeuArgGlnAsnLeuGlnAlaProValMetSerAsp 160
522 CGAGAAATGCCAAAAACAGAACAGAAAAAGCCACAGAAATCTTATGTGTAATTT 581
161 ArgGlnCysGlnLysThrGlnGlnGlyLysSerHisArgAsnSerLeuCysValLysPhe 180
582 GTAAGATTTTACCGCAATTTTGGGAGGTGGCGGTTGCTACTGCTACATGCAAGAC 641

Db 181 VallyValPheSerArgIlePheGlyGlyValAlaValAlaThrValIleCysLysAsp 200
Qy 642 AAGTCAGGAATTCAGAGTGGGCACTTCTATGGAGGGAGCGTGGCATCTACCAAT 701
Db 201 LysLeuGlnGlyIleGlyValAlaGlyHisPheMetGlyLysAspValGlyIleYrThrAsn 220
Qy 702 GTTACAAATATGATCTGATGAGAACACTGCTAAGGACAAG 746
Db 221 ValTyrLysTyrValSerTrpIleGlnAsnThrAlaLysAspLys 235

RESULT 4
US-10-246-006-2
; Sequence 2, Application US/10246006
; Publication No. US20030144493A1
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; TITLE OF INVENTION: Human Serine Protease
; FILE REFERENCE: 00-16
; CURRENT APPLICATION NUMBER: US/10/246,006
; CURRENT FILING DATE: 2002-09-17
; NUMBER OF SEQ. ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-246-006-2

Alignment Scores:
Pred. No.: 1,266-115 Length: 235
Score: 1247.00 Matches: 234
Percent Similarity: 99.57% Conservative: 0
Best Local Similarity: 99.57% Mismatches: 1
Query Match: 79.94% Indels: 0
DB: 12 Gaps: 0

US-10-037-270-482 (1-866) x US-10-246-006-2 (1-235)

Qy 42 ATGAAATATGCTCTTCTATTGGGTCCTGCTGGGACATTTTCTTGCTACTATCT 101
Db 1 MetLysTyrValPheTyrLeuGlyValLeuAlaGlyThrPhePhePheAlaAspSerSer 20
Qy 102 GTTCAGAAAGAACCCCTGCTCCCTATTGGTGTACCTCAAGTCTCACTCAACCCCTGT 161
Db 21 ValGlnLysGlnAspProAlaProTyrLeuValTyrLeuLysSerHisPheAsnProCys 40
Qy 162 GTGGGCTGCTCATCAAAACCCAGCTGGGTGCTGGCCCACTCACTCTATTACCAAT 221
Db 41 ValGlyValLeuIleLysProSerTrpValLeuAlaProAlaHisCysTyrLeuProAsn 60
Qy 222 CTGAAATGATGCTGGGAAATTTCAAGACAGACTCAGACGCTATCAAGACCAAT 281
Db 61 LeuLysValMetLeuGlyAsnPheLysSerArgValArgAspGlyThrGlnGlnThrIle 80
Qy 282 AACCCATTAGATGCTCGCTACTGGAACCTAGCTAGCGCCCAAGATGACCTC 341
Db 81 AsnProIleGlnIleValArgTyrTrpAsnTyrSerHisSerAlaProGlnAspAspLeu 100
Qy 342 ATGCTCATCAAGCTGCTAAGCCCTGCAATGCTCAATCCCAAGTCCAGCCCTTCCCTC 401
Db 101 MetLeuIleLysLeuAlaLysProAlaMetLeuAsnProLysValGlnProLeuThrLeu 120
Qy 402 GCCACCAACCAATGTCAGGCCAGGCACTGTCTTACTCTCAGTTTGAAGTGAACCA 461
Db 121 AlaThrThrAsnValArgProGlyThrValCysLeuLeuSerGlyLeuAspTrpSerGln 140
Qy 462 GAAACAGTGGCGCAGACCCCTGACTTGGCGAGAACCTGAGAGGCCCGCGATGTCAT 521
Db 141 GlnLysSerGlyArgHisProAspLeuArgGlnAsnLeuGlnAlaProValMetSerAsp 160
Qy 522 CGAAGATGCCAAAAAAGAAAGAAAGAAAGCAAGAAATTCCTATGTGTAATTT 581

Db 161 ArgGlyCysGlnLysThrGlnGlnGlyLysSerHisArgAsnSerLeuCysValLysPhe 180
Qy 582 GTCAAGTATTCAGCCGAATTTTGGGAGAGTGGCCGCTGCTCATCTGCAAGAC 641
Db 181 ValLysValPheSerArgIlePheGlyGlyValAlaValAlaThrValIleCysLysAsp 200
Qy 642 AAGTCAGGAATTCAGAGTGGGCACTTCTATGGAGGGAGCGTGGCATCTACCAAT 701
Db 201 LysLeuGlnGlyIleGlyValAlaGlyHisPheMetGlyLysAspValGlyIleYrThrAsn 220
Qy 702 GTTACAAATATGATCTGATGAGAACACTGCTAAGGACAAG 746
Db 221 ValTyrLysTyrValSerTrpIleGlnAsnThrAlaLysAspLys 235

RESULT 5
US-09-888-615-102
; Sequence 102, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOMMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CHENEDELL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ. ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 102
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-102

Alignment Scores:
Pred. No.: 4,096-113 Length: 265
Score: 1222.00 Matches: 234
Percent Similarity: 88.30% Conservative: 0
Best Local Similarity: 88.30% Mismatches: 1
Query Match: 78.33% Indels: 30
DB: 9 Gaps: 1

US-10-037-270-482 (1-866) x US-09-888-615-102 (1-265)

Qy 42 ATGAAATATGCTCTTCTATTGGGTCCTGCTGGGACATTTTCTTGCTACTATCT 101
Db 1 MetLysTyrValPheTyrLeuGlyValLeuAlaGlyThrPhePhePheAlaAspSerSer 20
Qy 102 GTTCAGAAAGAACCCCTGCTCCCTATTGGTGTACCTCAAGTCTCACTCAACCCCTGT 161
Db 21 ValGlnLysGlnAspProAlaProTyrLeuValTyrLeuLysSerHisPheAsnProCys 40
Qy 162 GTGGGCTGCTCATCAAAACCCAGCTGGGTGCTGGCCCACTCACTCTATTACCAAT 221
Db 41 ValGlyValLeuIleLysProSerTrpValLeuAlaProAlaHisCysTyrLeuProAsn 60
Qy 222 CTGAAATGATGCTGGGAAATTTCAAGACAGAGTCAAGCGCTCAAGACCAAT 281
Db 61 LeuLysValMetLeuGlyAsnPheLysSerArgValArgAspGlyThrGlnGlnThrIle 80
Qy 282 AACCCATTAGATGCTCGCTACTGGAACCTAGCTAGCGCCCAAGATGACCTC 341
Db 81 AsnProIleGlnIleValArgTyrTrpAsnTyrSerHisSerAlaProGlnAspAspLeu 100
Qy 342 ATGCTCATCAAGCTGCTAAGCCCTGCAATGCTCAATCCCAAGTCCAGCCCTTCCCTC 401
Db 101 MetLeuIleLysLeuAlaLysProAlaMetLeuAsnProLysValGlnProLeuThrLeu 120

OY	402	GCACACCAATGTACAGGCGAGCACCTGCTGTACTCTCAGAGTTGGACTGCACAA	461
Db	121	AlaThrThrAsnValArgProGlyThrValCysLeuLeuSerGlyLeuAspTrpSerGln	140
OY	462	GAAGAACGT-----	470
Db	141	GluAsnSerGlyLeuThrProGlnLeuGluProGlyHisLeuThrLeuHisArgGlyPro	160
OY	471	-----GGCGACACCTGACTTGGCGG	491
Db	161	AlaIleProAspTrpGlnArgHisAsnSerHisGluGlnGlyArgHisProAspLeuArg	180
OY	492	CAGAACCTGAGAGCCCCGTGATGTTCTGCAGAGATGCGCAAAAAACAGAACAGAAAA	551
Db	181	GlnAsnLeuGluAlaProValMetSerAspArgGluCysGlnIleThrGluGlnGlyArg	200
OY	552	AGCCACAGGAATTCCTTATGTGTGAATTTGTGAAAGTATTCAGCCGGAATTTTGGGGAG	611
Db	201	SerHisAspAsnSerLeuGlySerValLysProValLysValPheSerArgIlePheGlyGlu	220
OY	612	GTGGCCGTGCTCATCTGTCTGCAAAAGACAAGCTCCAGAGGATCGAGGGGGGCACTTC	671
Db	221	ValAlaValAlaThrValIleCysLysAspLysLeuGlnGlyIleGluValGlyHisPhe	240
OY	672	ATGGAGGGGAGCTGGCAGCTTACACCAATGTTTACAAATATGATCTCTGGATTGGAAC	731
Db	241	MetGlyGlyAspValGlyIleIleThrThrAsnValTyrLysTyrValSerTrpIleGluAsn	260
OY	732	ACTGCTAAGACAAG 746	
Db	261	ThrAlaLysAspLys 265	

RESULT 6

```

US-09-842-758-107
: Sequence 107, Application US/09842758
: Publication No. US20030083244A1
: GENERAL INFORMATION:
: APPLICANT: Vernet, Corine A. M.
: APPLICANT: Fernandes, Elma R
: APPLICANT: Gerlach, Valerie
: APPLICANT: Shimkets, Richard A
: APPLICANT: Malynkar, Uriel M
: APPLICANT: Boldog, Ferenc L
: APPLICANT: Zeyhuseen, Bryan D
: APPLICANT: Spytek, Kimberly A
: APPLICANT: Majumder, Kumud
: APPLICANT: Tchernev, Velizar T
: APPLICANT: Padigara, Muralidhara
: APPLICANT: Paturajan, Meera
: APPLICANT: Burgess, Catherine E
: APPLICANT: Gangolli, Esna A
: APPLICANT: Smithson, Glenda
: APPLICANT: Rastelli, Luca
: APPLICANT: MacDougall, John R
: APPLICANT: Taupier, Raymond J
: APPLICANT: Grose, William M
: APPLICANT: Edward, Szekeres S
: APPLICANT: Alsobrook II, John P
: TITLE OF INVENTION: No. US20030083244A1 Proteins and Nucleic Acids Encoding Same
: FILE REFERENCE: 15965-783
: CURRENT APPLICATION NUMBER: US/09/842,758
: CURRENT FILING DATE: 2001-04-25
: PRIOR APPLICATION NUMBER: 60/200,158
: PRIOR FILING DATE: 2000-04-26
: PRIOR APPLICATION NUMBER: 60/200,613
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,780
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/201,006
: PRIOR FILING DATE: 2000-05-01
: PRIOR APPLICATION NUMBER: 60/201,007
: PRIOR FILING DATE: 2000-05-01
: PRIOR APPLICATION NUMBER: 60/201,236

```

```

? PRIOR FILING DATE: 2000-05-01
? PRIOR APPLICATION NUMBER: 60/201,238
? PRIOR FILING DATE: 2000-05-01
? PRIOR APPLICATION NUMBER: 60/201,186
? PRIOR FILING DATE: 2000-05-02
? PRIOR APPLICATION NUMBER: 60/201,474
? PRIOR FILING DATE: 2000-05-03
? PRIOR APPLICATION NUMBER: 60/201,508
? PRIOR FILING DATE: 2000-05-03
? PRIOR APPLICATION NUMBER: 60/220,591
? PRIOR FILING DATE: 2000-07-25
? PRIOR APPLICATION NUMBER: 60/232,678
? PRIOR FILING DATE: 2000-09-15
? PRIOR APPLICATION NUMBER: 60/263,217
? PRIOR FILING DATE: 2001-01-22
? PRIOR APPLICATION NUMBER: 60/265,160
? PRIOR FILING DATE: 2001-01-30
? NUMBER OF SEQ ID NOS: 113
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 107
? LENGTH: 231
? TYPE: PRT
? ORGANISM: Homo sapiens
US-09-842-758-107

Alignment Scores:
Pred. No.:      1,24e-112
Score:          1217.00
Percent Similarity: 98.70%
Best Local Similarity: 98.70%
Query Match:    78.01%
                11
                11
Gaps:           0
                0

```

US-10-037-270-482 (1-866) X US-09-842-758-107 (1-231)

Qy	44	ATGAATATGCTCTCAATTTGGGTGCGCTGGGACATTTTCTTGCTGACATCT	102
Dd	1	MeLlytYValPhefYLeuolYValLeuolagLYthrPhePhePheAlaAspSer	20
Qy	102	GTTTCAGAAAGAACCCCTGCTCCCTATTGTGTGTAACCTCAAGTCTCAACCCCTGT	162
Dd	21	ValGlnLysGluAspProAlaProTyrLeuValTyrLeuLysSerAlaAspProCys	40
Qy	162	GTTGGGCTCTCATCAAAACCCAGCTGGGTGCTGGCCCCAGCTCACTGTATTACCAAT	222
Dd	41	ValGlyValLeuLLeuYProSerLTrpAlaLeuAlaProAlaHisCysTyrLeuProAsn	60
Qy	222	CTGAAGTATGTCGTGGGAAATTTTCAAGACAGAGTCAGACAGCGTATCGAAACAGAAATT	282
Dd	61	LeuLysValMetLeuGlyAsnPheLysSerLysArgValaLysAspGlyLThrGlnThrIle	80
Qy	282	AACCCGATTCAGATCGTCGGCTACTGGAATCAAGTCATAGCGGCCACAGATGACCTC	342
Dd	81	AsnProLlGlnLlLeValArgTyrTrpAsnTyrSerHisSerAlaProGlnAspAspLeu	100
Qy	342	ATGCTATATTAACCTGGCTTAAGCTCCCATGTCTCAATCCCAAGTCCAGCCCTTCCCTC	402
Dd	101	MetLeuLlLeuLysLeuLlAsyProAlaMetLeuAsnProLysValGlnProLeuThrLeu	120
Qy	402	GCCACACCAAAATGTCAGGCGACGACTCTCTGTACTACTCAGGTTTGACTGAGACCA	462
Dd	121	AlaThrThrAsnValLArgProGlyLThrValCysLeuLeuSerGlyLeuAspTrpSerGln	140
Qy	462	GAATAACAGTGGCCGACACCTCTGACTTGGCGGAACTGGAGGCCCCCGTATGTCGAT	522
Dd	141	GluAsnSerGlyArgHisProAspPheuArgGlnAsnLeuGlnAlaProValMetSerAsp	160
Qy	522	CGAGATATGCA-AAAAACAGAACAGAGAAAAACACAGAAATTCCTTAATGTGAATT	580
Dd	161	ArgGlnCysGlnLysAsnArgThrArgLysLysProGlnGluPheLeuMetCysGlnIle	180
Qy	581	TGTGAAGAATTCACCGCAATTTTGGGAGGTGGCGGTTCATCTCATCTGCAAGA	640

Db 181 CysGluSerIleGlnProAenPheTrpGlyGlyAlaArgCysTyrCysHisLeuGlnArg 200
Qy 641 CAAGCTCCAGGGAATCCAGGTGGGCACTTCATGGAGGAGGACCTCGGACATCTACACCAA 700
Db 201 GlnAlaProGlyAenAaGlyGlyAlaLeuHisGlyArgGlyAlaGlnHisLeuHisGln 220
Qy 701 TGTTTACAAATATGTATCTCTGGATTGAGA 730
Db 221 CysLeuGlnIleCysIleLeuApsSerGlu 230
RESULT 7
US-09-842-758-54
; Sequence 54, Application US/09842758
; Publication No. US20030083244A1
; GENERAL INFORMATION:
; APPLICANT: Verneet, Corine A. M.
; APPLICANT: Fernandez, Elma R.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Boidog, Ferenc L.
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Paturajan, Meera
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gangolli, Beha A.
; APPLICANT: Smithson, Glenda
; APPLICANT: Raetelli, Luca
; APPLICANT: MacDougall, John R.
; APPLICANT: Taupier, Raymond J.
; APPLICANT: Grose, William M.
; APPLICANT: Edwards, Szekeres S.
; APPLICANT: Alsobrook II, John P.
; TITLE OF INVENTION: No. US20030083244A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-783
; CURRENT APPLICATION NUMBER: US/09/842,758
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/200,158
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/200,613
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,780
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/201,006
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,007
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,236
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,186
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: 60/201,474
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/201,508
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/220,591
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/232,678
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 60/263,217
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 60/265,160
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 54
; LENGTH: 248
; TYPE: PRT

; ORGANISM: Gallus gallus
US-09-842-758-54
Alignment Scores:
Pred. No.: 8e-29 Length: 248
Score: 377.50 Matches: 88
Percent Similarity: 53.06% Conservative: 42
Best Local Similarity: 35.92% Mismatches: 100
Query Match: 24.20% Indels: 15
DB: 11 Gaps: 4
US-10-037-270-482 (1-866) x US-09-842-758-54 (1-248)
Qy 42 ATGAATAT-----GCTTCTATTGGGTGCTCGTGGACATTTTCTTGGCT 92
Db 1 MetLysPheLeuValLeuValAlaPheLeuGlyVal---AlaValAlaPheProIleSer 19
Qy 93 GACTCATCTGTTAGAAA-----GAAACCTGCTCCCTAT 128
Db 20 AspGluAspAspAspLysIleValGlyGlyTyrSerCysAlaArgSerAlaIleProTyr 39
Qy 129 TTGGTGTACTCAAGTCTCACTCAACCCCTGTTGGGCGTCTCATCAACCCAGCTGG 188
Db 40 GlnValSerLeuAsnSerGlyTyrHisPheCysGlyGlySerLeuIleSerSerGlnTyr 59
Qy 189 GTCTGGCCCAAGCTCACTGCTATTATCCAAATCTGAAGTGAATGCTGGAAATTTCAAG 248
Db 60 ValLeuSerAlaIleHisCysTyrLysSerSerIleGlnValLysLeuGlyGlyTyrAen 79
Qy 249 AGCAGAGTCAGAGCGGTACTGAACGACATTAACCCATTAGATGCTCCGCTACTGG 308
Db 80 LeuAlaIleGlnAspLysSerGlnGlnThrIleSerSerSerValIleArgHisSer 99
Qy 309 AACTACAGTATGATGCCCCACAGATGACCTCATCTATCAAGCTGAGTGAAGCTGCC 368
Db 100 GlyTyrAsnSerAsnThrLeuAsnAspIleMetLeuIleLysLeuSerLysAlaAla 119
Qy 369 ATCTCAATCCCAAGATCCAGGCCCTTCCCTCGCCACCAACCAATGTCAAGCCAGCACT 428
Db 120 ThrLeuAsnSerTyrValAsnThrValProLeuProThrSerCysValThrAlaGlyThr 139
Qy 429 GTCTGTCTACTCTCAGGTTTGACCTGAGCCAAAGAAACAGTGGCCGACACCTGACTTG 488
Db 140 ThrCysLeuIleSerGlyTyrGlyAsnThrLeuSerSerGlySerLeuTyrProAspVal 159
Qy 489 CGGCAGAACCTGAGGCCCCCGATGCTGATCGAAGTGCAGAAAGTATTCAGCCGAATTTTGGG 548
Db 160 LeuGlnCysLeuAsnAlaProValLeuSerSerSerGlnCysSerSerAlaTyrProGly 179
Qy 549 AAAGCCACAGAAATCTTATGTGAAATTTGTGAAGTATTCAGCCGAATTTTGGG 608
Db 180 ArgIleThrSerAsnMetIleCysIleGlyTyrLeuAsnGlyGlyLysAspSerCysGln 199
Qy 609 GAGGTGCGCTGTCTACTGATCTGCAAGACAGACCTCCAGGAATC-----GAG 659
Db 200 GlyAspSerGlyGlyProValValCysAsnGlyGlnLeuGlnGlyPheValSerTrpGly 219
Qy 660 GTGGGCACTTCATGGAGGAGCGTCGCATTAACCAATGTTAAATATGATATC 719
Db 220 IleGlyCysAlaGlnLysGlyTyrProGlyValTyrThrLysValCysAsnTyrValSer 239
Qy 720 TGGATTGAGAACT 734
Db 240 TrpIleLysThrThr 244
RESULT 8
US-09-842-758-53
; Sequence 53, Application US/09842758
; Publication No. US20030083244A1
; GENERAL INFORMATION:
; APPLICANT: Verneet, Corine A. M.
; APPLICANT: Fernandez, Elma R.
; APPLICANT: Gerlach, Valerie

APPLICANT: Shimkets, Richard A
APPLICANT: Malyanekar, Uriel M
APPLICANT: Boldog, Ferenc L
APPLICANT: Zerhusen, Bryan D
APPLICANT: Spytek, Kimberly A
APPLICANT: Majumder, Kumud
APPLICANT: Tchernev, Velizar T
APPLICANT: Padigar, Muralidhara
APPLICANT: Paturajan, Meera
APPLICANT: Burgess, Catherine E
APPLICANT: Gangoli, Bha A
APPLICANT: Smithson, Glenda
APPLICANT: Rastelli, Luca
APPLICANT: MacDougall, John R
APPLICANT: Taupier, Raymond J
APPLICANT: Groese, William M
APPLICANT: Edward, Szekeres S
APPLICANT: Alsobrook II, John P
FILE REFERENCE: 15966-783
FILE REFERENCE: No. US20030083244A1e1 Proteins and Nucleic Acids Encoding Same
CURRENT FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: 60/200,158
PRIOR FILING DATE: 2000-04-26
PRIOR APPLICATION NUMBER: 60/200,613
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,780
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/201,006
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,007
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,236
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,238
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,186
PRIOR FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: 60/201,474
PRIOR FILING DATE: 2000-05-03
PRIOR APPLICATION NUMBER: 60/201,508
PRIOR FILING DATE: 2000-05-03
PRIOR APPLICATION NUMBER: 60/220,591
PRIOR FILING DATE: 2000-07-25
PRIOR APPLICATION NUMBER: 60/222,678
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 60/263,217
PRIOR FILING DATE: 2001-01-22
PRIOR APPLICATION NUMBER: 60/265,160
PRIOR FILING DATE: 2001-01-30
NUMBER OF SEQ ID NOS: 113
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 53
LENGTH: 248
TYPE: PRT
ORGANISM: Gallus gallus
US-09-842-758-53
Alignment Scores:
Pred. No.: 1,27e-28 Length: 248
Score: 375.50 Matches: 87
Percent Similarity: 53.06% Conservative: 43
Best Local Similarity: 35.51% Mismatches: 100
Query Match: 24.07% Indels: 15
DB: 11 Gaps: 4
US-10-037-270-482 (1-866) x US-09-842-758-53 (1-248)
QY 42 ATCAATAT-----GCTTCTATTGGGTGCTCTGCGACATTTTCTTGTCT 92
DB 1 MetLyspHeuValIeuValAlaIleValGlyValThrVal---AlaIleProIleSer 19
QY 93 GACTCATCTGTTGAGAAA-----GAAGACCTGCTCCTAT 128

DB 20 AaPGLuAspAspAspLysLeValGlyGlyTyrSerCysAlaArgSerAlaIleProTyr 39
QY 129 TTGGTGTACTACCTGACCTGCTACCTTCAACCCCTGTGGGGTCTCTATCAACCCAGCTG 188
DB 40 GlnValSerLeuAsnSerGlyTyrHisPheCysGlyGlySerLeuIleSerSerGlnTTP 59
QY 189 GTGCTGGCCCCAGCTCACTGCTATTATTTCAATCTGAAGTGTGCTGGAAATTTCAAG 248
DB 60 ValLeuSerAlaAlaHisCysTyrTyrSerSerIleGlnValLeuGlyGlyTyrAsn 79
QY 249 AGCAGAGTCAGAGCCGTAAGTGAACAGACATTTAACCCATTGATCGTCCGCTAGTGG 308
DB 80 LeuAlaIleGlnAspLysSerGlnGlnThrIleSerSerSerValIleArgHisSer 99
QY 309 AACTACAGTCATAGCCCCCAAGATGACATGATCTATCAAGTGTGCTGACCTGCC 368
DB 100 GlyTyrAsnAlaAsnThrLeuAsnAspIleMetLeuIleLysLeuSerLysAlaAla 119
QY 369 ATGCTCATCCCAAGTCCAGCCCTTCCCTGCGACCAACATGCTCAGGCCAGCTGACT 428
DB 120 ThrLeuAsnSerTyrValAsnThrValProLeuProThrSerCysValThrAlaGlyThr 139
QY 429 GTCTGTCTACTCTCAGGTTTGAATGAGCCCAAGAAAGTGGCCGACACCTGACTTG 488
DB 140 ThrCysLeuIleSerGlyTyrGlyAsnThrLeuSerSerGlySerLeuTyrProAspVal 159
QY 489 CCGCAGAACCTGAGAGCCCTGATGCTGTGATTCAGAAATGCCAAAACAGAACAGAA 548
DB 160 LeuGlnCysLeuAsnAlaProValLeuSerSerSerGlnCysSerSerAlaTyrProGly 179
QY 549 AAAGCCACAGAAATCTTATGTGTGAATTTGTGAAGTATTCAGCCGAATTTTGGG 608
DB 180 ArgIleThrSerAsnMetIleCysIleGlyTyrLeuAsnGlyGlyLysAspSerCysGln 199
QY 609 GAGTGGCCGTTGCTCATCTGATCTGCAAAAGCAACCTCAGGAATC-----GAG 659
DB 200 GlyAspSerGlyGlyProValValCysAsnGlyGlnLeuGlnGlyIleValSerTpgly 219
QY 660 GTGGGCACTTATGAGAGGAGGACGTCGATCTACCAATGTTCAATATGATATC 719
DB 220 IleGlyCysAlaGlnLysGlyTyrProGlyValTyrThrLysValCysAsnTyrValSer 239
QY 720 TGGATTGAGACACT 734
DB 240 TrpIleLysThrThr 244
RESULT 9
US-09-842-758-55
Sequence 55, Application US/09842758
Publication No. US20030083244A1
GENERAL INFORMATION:
APPLICANT: Vernet, Corine A. M.
APPLICANT: Fernandes, Elma R.
APPLICANT: Gerlach, Valerie
APPLICANT: Shimkets, Richard A
APPLICANT: Malyanekar, Uriel M
APPLICANT: Boldog, Ferenc L
APPLICANT: Zerhusen, Bryan D
APPLICANT: Spytek, Kimberly A
APPLICANT: Majumder, Kumud
APPLICANT: Tchernev, Velizar T
APPLICANT: Padigar, Muralidhara
APPLICANT: Paturajan, Meera
APPLICANT: Burgess, Catherine E
APPLICANT: Gangoli, Bha A
APPLICANT: Smithson, Glenda
APPLICANT: Rastelli, Luca
APPLICANT: MacDougall, John R
APPLICANT: Taupier, Raymond J
APPLICANT: Groese, William M
APPLICANT: Edward, Szekeres S
APPLICANT: Alsobrook II, John P

FILE OF INVENTION: No. US-2003008344A1
FILE REFERENCE: 15966-783
CURRENT APPLICATION NUMBER: US/09/842,758
PRIOR APPLICATION NUMBER: 2001-04-25
PRIOR FILING DATE: 2000-04-26
PRIOR APPLICATION NUMBER: 60/200,613
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,780
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/201,006
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,007
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,236
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,238
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,186
PRIOR FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: 60/201,474
PRIOR FILING DATE: 2000-05-03
PRIOR APPLICATION NUMBER: 60/201,508
PRIOR FILING DATE: 2000-05-03
PRIOR APPLICATION NUMBER: 60/220,591
PRIOR FILING DATE: 2000-07-25
PRIOR APPLICATION NUMBER: 60/232,678
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 60/263,217
PRIOR FILING DATE: 2001-01-22
PRIOR APPLICATION NUMBER: 60/265,160
PRIOR FILING DATE: 2001-01-30
NUMBER OF SEQ ID NOS: 113
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 55
LENGTH: 246
TYPE: PRT
ORGANISM: Mus musculus
US-09-842-758-55

[illegible]

Oy	468	AGTGGCCCAACCCCTGACTTTGGCGCAGAACTCGGAAGCCCCCGGTATGTCTGATCCAGAA	527
Dd	151	GlyValAsnAsnProAspPheUenIncysValAspAlaProValLeuProGlnAlaAsp	170
Oy	528	TGCCAAAAACAAGAACAAGAAAACCACAGGAATTCTTATGTGTCAAATTTTGAAA	587
Dd	171	CysGluAlaSerTyrProGlyAspIleThrAsnAsnMetIleCysValIdjPheLeuGlu	190
Oy	588	GTATTACAGCCCAATTTTGGGAGAGGTGCCTGTGTACTGTATCTGCACAAAGCAAGCTC	647
Dd	191	GlyGlyLysAspSerCysGlnGlyLysAspSerGlyGlyProValCysAsnGlyGluLeu	210
Oy	648	CAGGGAAATC-----GAGGTGGGCACTTCATGGGAGGGAGAGCTCGGCATCTACACC	698
Dd	211	GlnGlyIleValIserTripglyTyrIleCysAlaGlnProAspAlaProIleValYalYthr	230
Oy	699	AATGTTCACAATAATGTATCTCGAATGGAACACTGCTAACAGAC	743
Dd	231	LysValCysAsnGlyValAspTripleGlnAsnThrIleAlaAsp	245

```

RESULT 10
US-09-923-779-154
: Sequence 154, Application US/05923779
: Patent No. US20020076721a1
: GENERAL INFORMATION:
: APPLICANT: Pyle, Ruth A.
: APPLICANT: Xu, Jiangchun
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
: FILE REFERENCE: 210121.553
: CURRENT APPLICATION NUMBER: US/09/923,779
: CURRENT FILING DATE: 2001-08-06
: NUMBER OF SEQ ID NOS: 155
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 154
: LENGTH: 247
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-923-779-154

```

Pred. No.:	1,576-24	247
Score:	334.50	75
Percent Similarity:	52.838	Conservative: 37
Best Local Similarity:	35.384	Mismatch: 97
Query Match:	21.444	Indels: 3
DB:	9	Gaps: 1

US-10-037-270-482 (1-866) x US-09-923-779-154 (1-247)

Qy	108	AAAGAGACCCCTGCTCCCTATTGGTGTACTCAAGTCTCAACCCCTGTGGGC	167
		:::	:::
Db	31	GIUGIuAnSerValProtyrGlnValSerLeuAnSerGlyTrnHisPheGyGly	50
		:::	:::
Qy	168	GTCCTCATAAACCAAGCTGGTGTGGCCCCAGCTACTGCTCTTTTACCAATCTGAA	227
		:::	:::
Db	51	SeLeuHisSerGlnGlnTrpValValSerLagLynHisCylXylSerArgLleGln	70
		:::	:::
Qy	228	GTCATCTGGGAAATTTCAAGACGAGCTCAGACGCTACTGAAACAGACAAATACCC	287
		:::	:::
Db	71	ValArgLeuGlyGlnHisAsnLecIValValLeuGlnGlyAnGlnGlnPheIleAsnAla	90
		:::	:::
Qy	288	ATTCAATGCTCCGCTACTGGAATACAGTACGATATAGCGCCACAGATGACTCATGCTC	347
		:::	:::
Db	91	AlaValIleIleArgHisProLysTrpGlyAnSerArgTrnHisLeuAspAnAspIleLeuLeu	110
		:::	:::
Qy	348	ATTCAGCTGCTAAAGCTGCGACTGCTCAATCCCAAGTCCAGCCCTTCCCTGGCCACC	407
		:::	:::
Db	111	IleLysLeuSerSerProLValIleLeuSerArgValSerAlaIleSerLeuProTrn	130
		:::	:::
Qy	408	ACCAATGTAGCCAGGCACTGTCTGTACTTTCAGGTTTGGACTGGAGCCAAAGAAC	467
		:::	:::


```

Db      111 AlaProAlaAlaGlyThrGluSerLeuIleSerGlyTyrGlyAsnThrLeuSerSer 150
QY      468 AGTGGCCGACACCTGACCTGTCGGGAGAACCTGGAGACCCCGCTGATGCTGATGACGAA 527
Db      151 G1yAlaAspTyrProAspGluLeuGlnCysLeuAspAlaProValLeuSerGlnAlaGlu 170
QY      528 TSCCAAAAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 587
Db      171 CysGluAlaSerTyrProGlyLysIleThrAsnAsnMetPheCysValGlyPheLeuGlu 190
QY      588 GTATTCAAGCGAATTTTGGGAGGTGGCGCTGCTACTGATCTGATCTGCAAGCAAGCTC 647
Db      191 GlyGlyLysAspSerCysGlnGlyAspSerGlyGlyProValValSerAsnGlyLeuLeu 210
QY      648 CAGGGAATC-----GAGTGGGGGCACTTCATGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 698
Db      211 GlnGlyIleValAspSerTyrGlyTyrGlyCysAlaGlnLysAsnArgProGlyValTyrThr 230
QY      699 AATGTTTCAAAATATGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 734
Db      231 LysValTyrAsnTyrValAspTyrIleLysAspThr 242

```

RESULT 11

US-09-898-837A-50
Sequence 50, Application US/09898837A
Publication No. US2003007697A1

GENERAL INFORMATION:

APPLICANT: Quinn, Kerry E.
APPLICANT: Spytek, Kimberly A.
APPLICANT: Majumder, Kumud
APPLICANT: Vernet, Corine
APPLICANT: Herrmann, John L.
APPLICANT: Burgess, Catherine
APPLICANT: Fernandes, Elma
APPLICANT: Taupier Jr., Raymond
APPLICANT: Rastelli, Luca
APPLICANT: Curagen Corporation
APPLICANT: Gerlach, Valerie L.
APPLICANT: MacDougall, John R.
TITLE OF INVENTION: NOVEL SERINE/THREONINE PROTEIN-KINASE LIKE PROTEINS AND
FILE REFERENCE: 15966-598 CIP
CURRENT APPLICATION NUMBER: US/09/898,837A
CURRENT FILING DATE: 2001-07-03
PRIOR APPLICATION NUMBER: U.S.S.N. 60/165,986
PRIOR FILING DATE: 1999-11-17
PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,839
PRIOR FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: U.S.S.N. 60/195,637
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: U.S.S.N. 60/197,080
PRIOR FILING DATE: 2000-04-13
PRIOR APPLICATION NUMBER: U.S.S.N. 60/232,677
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: U.S.S.N. 60/181,347
PRIOR FILING DATE: 2000-02-09
PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,195
PRIOR FILING DATE: 2000-04-03
PRIOR APPLICATION NUMBER: U.S.S.N. 60/215,906
PRIOR FILING DATE: 2000-07-03
PRIOR APPLICATION NUMBER: U.S.S.N. 09/715,427
PRIOR FILING DATE: 2000-11-16
NUMBER OF SEQ ID NOS: 53
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 50
LENGTH: 215
TYPE: PRT
ORGANISM: Homo sapiens
US-09-898-837A-50

Alignment Scores: 6.68e-24 Length: 215
Pred. No.:

Score: 328.00 Matches: 79
Percent Similarity: 52.91% Conservative: 39
Best Local Similarity: 35.43% Mismatches: 93
Query Match: 21.03% Indels: 12
DB: 11 Gaps: 4

US-10-037-270-482 (1-866) x US-09-898-837A-50 (1-215)

```

QY      69 CCGCGTGGGACATTTTCTTCTGACATCATCTGTCAGAAAGAACCCCTGCTCCCTAT 128
Db      1 ILeValGlyGlyTyrThrCysAlaAlaAsnSerIle-----ProTyr 14
QY      129 TTGGTGTACTCTCAAGTCT---CACTTCAACCCCTGTGTGGGCGCTCTCATCAAGCCAC 185
Db      15 GlnValSerLeuAsnSerGlyHisPhe-----CysGlyGlySerLeuIleAsnSerGln 32
QY      186 TTGGTGTCTGGCCCCAGCTCAGTCTGATTTTACCAAAATCTGAAGTGTGCGAAATTTTC 245
Db      33 TPrValValSerAlaAlaHisCysTyrLysSerArgIleGlnValArgLeuGlyGluHis 52
QY      246 AAGACACAGTGCAGAGCGGTACTGACACAGCAATTAACCCATTCAGATCGCGCTAC 305
Db      53 AenIleAspValLeuGlnGlyAsnGlnInPheIleAsnAlaAlaLysIleIleThrHis 72
QY      306 TGGAACTACAGTCATAGCGCCCCACAGATGACCTCATCTCATCAAGCTGCTAAGCT 365
Db      73 ProAsnPheAsnGlyAsnThrLeuAsnAsnIleMetLeuIleLysLeuSerSerPro 92
QY      366 GCCATGCTCAATCCCAAGTCCAGCCCTTCCCTCGCCACCAACCAATCTCAGGCCAGC 425
Db      93 AlaThrLeuAsnSerArgValAlaThrValSerLeuProArgSerCysAlaAlaGly 112
QY      426 ACTGTCGTCTACTCTCTCAGCTTTGAGCTGAGCCCAAGAAACAGTGGCCGACACCTGAC 485
Db      113 ThrGlnCysLeuIleSerGlyTyrGlyAsnThrLysSerSerGlySerSerTyrProSer 132
QY      486 TTGGCGCAGAACCTGAGGCGCCCGTGAATGTGATTCGAGATGCCAAAGAACAGAACAA 545
Db      133 LeuLeuGlnCysLeuLysAlaProValValCysAsnGlyGlnLeuGlnGlyIleValSerTyr 152
QY      546 GGAAGAACCCACAGAAATCTTATGTGTGAATTTGTGAAGATTCAGCCGAATTTT 605
Db      153 GlyGlnIleThrGlyAsnMetIleCysValGlyPheLeuGlnGlyLysAspSerCys 172
QY      606 GGGAGAGTGGCCGCTGCTACTGATCTGATCTGCAAGCAAGCAAGCTCAGGAGATC----- 656
Db      173 GlnGlyAspSerGlyProValValCysAsnGlyGlnLeuGlnGlyIleValSerTyr 192
QY      657 GAGTGGGCGCACTTCATGAGGAGGAGCGTGGCATCTACACCAATGTTTACAAATATGTA 716
Db      193 GlyTyrGlyCysAlaGlnLysAsnLysProGlyValTyrThrLysValCysAsnTyrVal 212
QY      717 TCCTGGATT 725
Db      213 AsnTyrPile 215

```

RESULT 12

US-10-021-368-7
Sequence 7, Application US/10021368
Publication No. US20020106367A1

GENERAL INFORMATION:

APPLICANT: Band, Wima
TITLE OF INVENTION: NES-1 POLYPEPTIDES, DNA, AND RELATED
MOLECULES AND METHODS

NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804

COMPUTER READABLE FORM:


```

Qy      423  GGCATGTCGTCTCTACTCTCTCAGGTTTGACGTGGAGCCAGAAAGAAAGTGGCCGACACCT 482
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      113  GLythrgInGcYsleuileSerelyTTPGlyAenThrLysSerSerelyThrseryrPro 132
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy      483  GACTTCCGCGACAGACCTGGAGGCCCCCGTGATGTCTGATCGAGATCGCAAAAAACGAA 542
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      153  ProGlyGlnIleThrsSeramhetPheCysAlaGlyTrleuGlnIleGlyLysAepSer 172
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      603  TTTGGGGAGGTGGCCGTTTCTACTGTCACTGCACTCGCAAGACAGACGTCGAGGAATC----- 656
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      173  CysGlnGlyAepSerSerelyGlyLysProValValCysSerSerelyLysleuGlnIleValSer 192
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy      657  --GAGGTGGGGCACTTCATGAGGAGGGAGCGTCGCGCATCTACCAATGTTTCAAAATAT 713
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      193  TrpGlySerGlyCysAlaGlnLysAenLysPProGlyValIlyrThrsYsValCysAenTyr 212
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
US-10-239-663-39
; Sequence 39, Application US/10239663
; Publication No. US20030139572A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia, K.
; APPLICANT: Smith, Randall, F.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kabnick, Karen
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GPS0018
; CURRENT APPLICATION NUMBER: US/10/239,663
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09226
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/1192,158
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,668
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/200,166
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-239-663-39

Alignment Scores:
Pred. No.: 3 45e-22 Length: 243
Score: 311.00 Matches: 85
Percent Similarity: 50.81% Conservative: 41
Best Local Similarity: 34.27% Mismatches: 104
Query Match: 19.94% Indels: 18
DB: 12 Gaps: 7

US-10-037-270-482 (1-866) x US-10-239-663-39 (1-243)

Qy      27  AAGACCTGCATCCATGAAATATGTCCTTCTATTTGGGTGTCCTCGCTGGACATTTTTC 86
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      4  LysSerTrpAenPheLeuSerMetLeuAenPheProValAlaLeuAaA-----Phe 20
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy      87  TTTCGTCATCTATCTGTTTCAGAAAGAACCCGTCGCCCATTTTGGTGATCACTCAAGTC 146
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      21  AenPProAepTyrThrVal-----SerSerThrProProTyrTrleuValIlyrLeuYsSer 38
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Oy 147 CACTTCAACCCGCTGTGGCGCGCTTCATCAAAACCCAGCGGGTGTGGCCCGCAAGCTCAAG
    ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 39 AepTyrLeuProCysAlaGlyValIleuIleHisProLeuTrpValIleThrAlaIleHis
    ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 207 TGTCTATTACCAATCTGAAAGTGAATGCTGGGAAATTTCAAGAGCAGACTCAGACGGT
    ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 59 CysAsnLeuProIlysLeuArgValIleLeuGly-----ValThrIleProAlaAspSer
    ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 267 ACTGA-----CAGACAATTMAOCCATTCAATATGCTCCGCTACATGCACTACAGT
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 77 AsnGluIysHisIleuGlnValIleGlyTyrGluIysMetLeuIleHisIleProHisPheSer
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 318 CATAGCCGCCACAGATGACCTCATGCTCATCAAGCTGGAGCTGAAGCTCGACTGCTCAAT
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 97 ValThrSerIleLeuPheHisAspIleMetLeuIleIleValLeuIysThrGluAlaGluLeuAsn
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 378 CCCAAAGTCGACGCCCTTCCTCCCTGGCCACCAACCAATGTCAGGCCAGGCACTGTCTGT
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 117 AspTyrValIysIleuAlaAsnLeuProTyrGlnThrIleSerGluAsnThrMetCysSer
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 438 CTCTCAGGTTTGGACTGTGAGACCCAAAGAAACAGTGGCCGACACCTTCATCTGGCGCAAG
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 137 ValSerThrTrpSerTyrAsnValCysAspIleTyrGlySerGluProAspSerLeuGlnThr
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 498 CTGAGAGCCCCCGTGATGTCTGATTCGAAATGCGCAAAAACAAACAAGAAAGAAACCAAC
    :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 157 ValAsnIleSerValIleSerIysProGlnCysArg-----AspAlaTyrIysThrTyr
    :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 558 -----AGAAATTCCTTATGTGTGAATTTGTGAAGATATTCAGCCGAATTTTGGG
    :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 175 AsnIleThrGluAsnMetLeuCysValGlyIleValProGlyArgGlnProCysIys
    :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 609 GAGGTGGCCGCTGCTACTGCTCATCTGCAAGAAACAACCTCCAGGAGATC-----GAGG
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 195 GluValSerAlaIleProAlaIleCysAsnGlyMetLeuGlnIleIleuSerPheAla
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 660 GTGGGCACTTCATGGAGGAGGAGCGTGGCATCTTACACCAATGTTTACAAATATGTATCC
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 215 AspGlySerValIleuArgAlaAspValGlyIleTyrAlaIysIlePheTyrIleProC
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 720 TGGATTGAGAACACTGCTAAGAC 743
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 235 TrpIleGluAsnValIleGlnAsn 242
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 15
US-10-137-870-248
; Sequence 248, Application US/10137870
; Publication No. US20030138883A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P330R1C155
; CURRENT APPLICATION NUMBER: US/10/137, 870
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See Palm or File Wrapper

```

; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 248
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-137-870-248

Alignment Scores:

Pred. No.:	1.37e-21	Length:	241
Score:	305.00	Matches:	83
Percent Similarity:	50.41%	Conservative:	41
Best Local Similarity:	33.74%	Mismatches:	104
Query Match:	19.55%	Indels:	18
DB:	12	Gaps:	7

US-10-037-270-482 (1-866) x US-10-137-870-248 (1-241)

```
QY 42 ATGAAATATGCTTC-----TATTGGGTGCTCGCTGGACATTTCTTGTCT 92
Db 1 MetLysPheIleLeuLeuTPAlaLeuLeuAsnLeuThrValAlaLeuAlaPheAsnPro 20
QY 93 GACTCATCTGTGTGAGAAAGAACCCCTGCTCCCTATTGGTGTACTCAAGTCTCACTTC 152
Db 21 AspTyrThrVal-----SerSerThrProProTyrLeuValTyrLeuLysSerAspTyr 38
QY 153 AACCCCTGTGTGGCGCTCTCATCAAAACCAGCTGGTGTGCTGGCCCACTCACTGCTAT 212
Db 39 LeuProCysAlaGlyValLeuIleHisProLeuTyrValIleThrAlaAlaHisCysAsn 58
QY 213 TTACCAATCTGAAAGTAGTGTGGGAAATTTCAAGACAGAGTCAGACGGTACTGAA 272
Db 59 LeuProLysLeuArgValIleLeuGly-----ValThrIleProAlaAspSerAsnGlu 76
QY 273 -----CAGCAATTAACCCCATTCAGATGCTCGCTCACTGAGACTACAGCATATAGC 323
Db 77 LysHisLeuGlnValIleGlyTyrGlnLysMetIleHisProHisPheSerValThr 96
QY 324 GCCCCACAGATGACCTCATGCTCATCAAGCTGAGCTAAGCTGACATGCTCAATCCCAA 383
Db 97 SerIleAspHisAspIleMetLeuIleLysLeuLysThrGlnAlaGlnLeuAsnAspTyr 116
QY 384 GTCCAGCCCTTCCCTCGCCACCAACCAATGTCAGGCCAGGCACTGTCTGTACTCTCA 443
Db 117 ValLysLeuAlaAsnLeuProTyrGlnThrIleSerGlnAsnThrMetCysSerValSer 136
QY 444 GTTTGGACTGAGCCAGAAACAGTGGCCGACACCTGACTTGCGGCAAGACTGGAG 503
Db 137 ThrTTrpSerTyrAsnValCysAspIleTyrLysGlnProAspSerLeuGlnThrValAsn 156
QY 504 GCCCCGCTGATGCTGATCGAGATGCCAAAAACAGAACAGAAAAAGCCAC----- 557
Db 157 IleSerValIleSerLysProGlnCysArg-----AspAlaTyrLysThrTyrAsnIle 174
QY 558 ---AGGAATTCCTTATGTGTGAATTTGTGAAGATATTCAACCGAATTTTGGGGAGTG 614
Db 175 ThrGlnAsnMetLeuCysValGlyIleValProGlyYargArgGlnProCysLysGlnVal 194
QY 615 GCGGTGCTACTGTCATCTGCAAAAGACAAAGCTCAGGGAATC-----GAGGTGGGG 665
Db 195 SerAlaIleProAlaIleCysAsnGlyMetLeuGlnGlyIleLeuSerPheAlaAspGly 214
QY 666 CACTTCATGGAGGAGGAGTGGGATCTTACCAACCAATGTATATATATCTGATGATT 725
Db 215 CysValLeuArgAlaAspValGlyIleTyrAlaLysIlePheTyrTyrIleProTyrIle 234
QY 726 GAGAAACACTGCTAAGAC 743
Db 235 GluAsnValIleGlnAsn 240
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2003, 23:05:46 ; Search time 3495 Seconds
(without alignments)
10136.691 Million cell updates/sec

Title: US-10-037-270-482
Perfect score: 866
Sequence: 1 ggcactactaccctcagctca.....tcggaatcgtacatcactagt 866

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813366 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenBank1:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pin:*
35: em_hcg_rod:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	862.8	99.6	1060	AX455324	AX455324 Sequence
2	857.8	99.1	890	AX442630	AX442630 Sequence
3	850.8	98.2	865	AR240728	AR240728 Sequence
4	850.8	98.2	865	AX318279	AX318279 Sequence
5	834	96.3	865	AR263885	AR263885 Sequence
6	730	84.3	730	AX350386	AX350386 Sequence
7	728.4	84.1	730	AX350388	AX350388 Sequence
8	706.4	81.6	708	AX247864	AX247864 Sequence
9	706.4	81.6	708	AX455326	AX455326 Sequence
10	703.8	81.3	721	AX350481	AX350481 Sequence
11	606.4	70.0	798	AX360087	AX360087 Sequence
12	527	60.9	1855	AK097648	AK097648 Sequence
13	511.2	59.0	705	AR240729	AR240729 Sequence
14	511.2	59.0	705	AX318281	AX318281 Sequence
15	359.8	41.5	479	BD113886	BD113886 EST and e
16	259	29.9	31047	AC093638	AC093638 Homo sapi
17	246	28.4	379	BD047802	BD047802 Sequence
18	196.8	22.7	136812	AC137710	AC137710 Mus muscu
19	196.8	22.7	204476	AC117613	AC117613 Mus muscu
20	193	22.3	225962	AC094695	AC094695 Rattus no
21	173.6	20.0	804	RNTRY1	V01273 Rat mRNA en
22	169.4	19.6	814	NMTTRYAR	X04574 Mouse mRNA
23	167.8	19.4	860	GGU15155	U15155 Gallus gall
24	167.8	19.4	864	AB009661	AB009661 Mus muscu
25	164.4	19.0	797	RNTRY2	V01274 Rat mRNA en
26	159.6	18.4	773	BTTRYAPAP	X54703 Bovine mRNA
27	155.4	17.9	805	4	BTTRYAPAP
28	154.4	17.8	741	9 HUMNA	D45417 Human mRNA
29	154.4	17.8	790	6 E15808	E15808 Human mRNA
30	154.4	17.8	821	9 BC030238	BC030238 Homo sapi
31	154.4	17.8	853	9 HSTRPIV	X72781 Homo sapien
32	154.4	17.8	875	9 BC030260	BC030260 Homo sapi
33	154	17.8	850	6 AX333266	AX333266 Sequence
34	154	17.8	850	9 HSTRYIVB	X71345 H. sapiens m
35	153.2	17.7	802	6 AX527621	AX527621 Sequence
36	153.2	17.7	802	9 HUMTRPSGNB	M27602 Human pancr
37	153.2	17.7	954	6 AX014870	AX014870 Sequence
38	152.8	17.6	741	6 E01617	E01617 cDNA encodi
39	152.8	17.6	744	6 E09633	E09633 DNA encodin
40	152.8	17.6	807	9 HSTRYIII	X15505 Human mRNA
41	145	16.7	825	4 BOVPCP	D38507 Bos taurus
42	145	16.7	825	6 AR103928	AR103928 Sequence
43	144.6	16.7	819	4 DGSTRYPA	M11589 Dog pancrea
44	142.8	16.5	860	5 GGU15157	U15157 Gallus gall
45	142	16.4	699	6 AX206893	AX206893 Sequence

ALIGNMENTS

RESULT 1
AX455324
LOCUS AX455324 1060 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 1 from Patent WO0206331.
ACCESSION AX455324
VERSION AX455324.1 GI:21714507
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Meyer, R.
TITLE 14087, a novel serine protease molecule and uses therefor
JOURNAL Patent: WO 0206331-A 1 24-JAN-2002;

Pred. No. is the number of results predicted by chance to have a

FEATURES
source
Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers
1.1060
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
164.871
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD38234.1"
/db_xref="GI:21714508"
translation="MKTVFYGLVLAGTFFPDDSSVQEDPAPLVYIKSHFPCVGL
IKPSWLAAPHCYLPNLKWLGNFKSRVRDQETNPIDQIVAYMYSHSAPQDDML
IKLAKPAMHPKQVPLPLATNNVRPGVCLSLGIDMSQENSHPLRLQLAEAPVMS
RECKTEOGKSHNSLCKVFKFVFSRIFGEVAATVYCKDKLGIEVGHMGSDVGIY
TNYKYVSWTEMTAKDK"

BASE COUNT 309 a 259 c 237 g 251 t 4 others
ORIGIN

Query Match 99.8%; Score 862.8; DB 6; Length 1060;
Best Local Similarity 99.8%; Pred. No. 2,7e-232;
Matches 864; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGCACCTACTCCCTGAGCTAAGGGGGAAGAGCTGGATCCACATGAATATGCTCTATT 60
DB 123 GACACTTACTCCCTGAGCTAAGGGGGAAGAGCTGGATCCACATGAATATGCTCTATT 182
QY 61 TGGGTGCTCCGCTGGGACATTTTCTTCTGCTGATCATCTGTCAGAAAGACCCCTG 120
DB 183 TGGGTGCTCCGCTGGGACATTTTCTTCTGCTGATCATCTGTCAGAAAGACCCCTG 242
QY 121 CTCCTATTGTTGTATCTCAAGTCTCACTTCAACCCCTGTGTGGCGTCTCATCAAC 180
DB 243 CTCCTATTGTTGTATCTCAAGTCTCACTTCAACCCCTGTGTGGCGTCTCATCAAC 302
QY 181 CCAAGCTGGGTGCTGGCCCCAGCTCACTGCTATTACCAAACTGAAAGTATGCTGGAA 240
DB 303 CCAAGCTGGGTGCTGGCCCCAGCTCACTGCTATTACCAAACTGAAAGTATGCTGGAA 362
QY 241 ATTTCAGAGCAGAGTCAGAGAGGTCATGACAGACAAATTAACCCATCAATGCTCC 300
DB 363 ATTTCAGAGCAGAGTCAGAGAGGTCATGACAGACAAATTAACCCATCAATGCTCC 422
QY 301 GCTACTGSACTACAGTCATAGGCCCCCAAGATGACTCATGCTCATCAAGCTGGCTA 360
DB 423 GCTACTGSACTACAGTCATAGGCCCCCAAGATGACTCATGCTCATCAAGCTGGCTA 482
QY 361 AGCTGTCATGCTCAATCCCAAGTCCAGCCCTTCCCTGGCCACCAATGTCAGGC 420
DB 483 AGCTGTCATGCTCAATCCCAAGTCCAGCCCTTCCCTGGCCACCAATGTCAGGC 542
QY 421 CAGGCACTGCTGCTCACTCAAGTCTCAAGTTCAGAGCCAAAGAAACAGTGGCCGAC 480
DB 543 CAGGCACTGCTGCTCACTCAAGTTCAGAGTTCAGAGCCAAAGAAACAGTGGCCGAC 602
QY 481 CTGACTTGGCGGAGAACTGGAAGCCCGCTGATGTCTGATCGAAGATGCCAAAAACAG 540
DB 603 CTGACTTGGCGGAGAACTGGAAGCCCGCTGATGTCTGATCGAAGATGCCAAAAACAG 662
QY 541 AACAGGAAAAAGCCACAGGAATTCCTTATGTGTGAATTTGTGAAGTATTCAGCGAA 600
DB 663 AACAGGAAAAAGCCACAGGAATTCCTTATGTGTGAATTTGTGAAGTATTCAGCGAA 722
QY 601 TTTTGGGAGAGTGGCGCTGCTGCTATGTCATGTCAGAAAGAGTTCAGAGG 660
DB 723 TTTTGGGAGAGTGGCGCTGCTGCTATGTCATGTCAGAAAGAGTTCAGAGG 782
QY 661 TGGGGACTTCAATGAGGAGGAGGAGTGGGATCTACACCAATGTTCAATATATATCTT 720
DB 783 TGGGGACTTCAATGAGGAGGAGGAGTGGGATCTACACCAATGTTCAATATATATCTT 842
QY 721 GGAATTGAGAACCTGCTAAGAGCAAGTGAAGCCCTACTTCCCTGCTGATTCAGCTGC 780
DB 842 GGAATTGAGAACCTGCTAAGAGCAAGTGAAGCCCTACTTCCCTGCTGATTCAGCTGC 842

DB 843 GGAATTGAGAACCTGCTAAGAGCAAGTGAAGCCCTACTTCCCTGCTGATTCAGCTGC 902
QY 781 TCTGCATGAGACTATCAAGCAGATATTTTCCCTGATTCATTAATTAATTCAGATG 840
DB 903 TCTGCATGAGACTATCAAGCAGATATTTTCCCTGATTCATTAATTAATTCAGATG 962
QY 841 AAAATTGGGAATGTAGCATCTAGT 866
DB 963 AAAATTGGGAATGTAGCATCTAGT 988

RESULT 2
AX342630 890 bp DNA linear PAT 12-JAN-2002
LOCUS Sequence 27 from Patent WO0198468.
DEFINITION AX342630
ACCESSION AX342630
VERSION AX342630.1 GI:18152027
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
1 Yue, H., Elliott, V. S., Gandhi, A. R., Lal, P., Au-Young, J.,
Tribouley, C. M., Deleage, A. M., Baughn, M. R., Nguyen, D. B., Lee, E. A.,
Hafalia, A., Khan, F. A., Wallis, N. K., Yao, M. G., Lu, D. A., Patterson, C.,
Tang, Y. T., Walsh, R. T., Azimzai, Y., Ramkumar, J., Xu, Y. and Reddy, R.,
Patent: WO 0198468-A 27-DEC-2001;
Incyte Genomics, Inc. (US)
Location/Qualifiers

FEATURES
source
1.890
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7474081CB1"

BASE COUNT 249 a 230 c 203 g 208 t
ORIGIN

Query Match 99.1%; Score 857.8; DB 6; Length 890;
Best Local Similarity 99.8%; Pred. No. 6.8e-231;
Matches 859; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGCACCTACTCCCTGAGCTAAGGGGGAAGAGCTGGATCCACATGAATATGCTCTATT 60
DB 22 GGCACCTACTCCCTGAGCTAAGGGGGAAGAGCTGGATCCACATGAATATGCTCTATT 81
QY 61 TGGGTGCTCCGCTGGGACATTTTCTTCTGCTGATCATCTGTCAGAAAGACCCCTG 120
DB 82 TGGGTGCTCCGCTGGGACATTTTCTTCTGCTGATCATCTGTCAGAAAGACCCCTG 141
QY 121 CTCCTATTGTTGTATCTCAAGTCTCACTTCAACCCCTGTGTGGCGTCTCATCAAC 180
DB 142 CTCCTATTGTTGTATCTCAAGTCTCACTTCAACCCCTGTGTGGCGTCTCATCAAC 201
QY 181 CCAAGCTGGGTGCTGGCCCCAGCTCACTGCTATTACCAATCTGAAAGTATGCTGGAA 240
DB 202 CCAAGCTGGGTGCTGGCCCCAGCTCACTGCTATTACCAATCTGAAAGTATGCTGGAA 261
QY 241 ATTTCAGAGCAGAGTCAGAGAGGTCATGACAGACAAATTAACCCATTCAGATGCTCC 300
DB 262 ATTTCAGAGCAGAGTCAGAGAGGTCATGACAGACAAATTAACCCATTCAGATGCTCC 321
QY 301 GCTACTGSACTACAGTCATAGGCCCCCAAGATGACTCATGCTCATCAAGCTGGCTA 360
DB 322 GCTACTGSACTACAGTCATAGGCCCCCAAGATGACTCATGCTCATCAAGCTGGCTA 381
QY 361 AGCTGTCATGCTCAATCCCAAGTCCAGCCCTTCCCTGGCCACCAATGTCAGGC 420
DB 382 AGCTGTCATGCTCAATCCCAAGTCCAGCCCTTCCCTGGCCACCAATGTCAGGC 441
QY 421 CAGGCACTGCTGCTCACTCAAGTTCAGAGTTCAGAGCCAAAGAAACAGTGGCCGAC 480
DB 442 CAGGCACTGCTGCTCACTCAAGTTCAGAGTTCAGAGCCAAAGAAACAGTGGCCGAC 501

QY 481 CTGACTTGGGCGAGAACTTGAGAGGCCCCGCTGATGTCGAGATGCGCAAAAAAG 540
 DB 502 CTGACTTGGGCGAGAACTTGAGAGGCCCCGCTGATGTCGAGATGCGCAAAAAAG 561
 QY 541 AACAGAGAAAAAGCCACAGAGAAATTCCTTATGTTGTAATTTGTAAGTAATTCAGCCGA 600
 DB 562 AACAGAGAAAAAGCCACAGAGAAATTCCTTATGTTGTAATTTGTAAGTAATTCAGCCGA 621
 QY 601 TTTTGGGAGAGGTGGCCGTTGCTACTGTCATCTGCAAGACAGCTCCAGGGAATGAGG 660
 DB 622 TTTTGGGAGAGGTGGCCGTTGCTACTGTCATCTGCAAGACAGCTCCAGGGAATGAGG 681
 QY 661 TGGGCGACTTCATGAGAGGGGAGCTGCGGACTTACACCAATGTTTCAAAATATGATTCCT 720
 DB 682 TGGGCGACTTCATGAGAGGGGAGCTGCGGACTTACACCAATGTTTCAAAATATGATTCCT 741
 QY 721 GGATTGAGAACACTGCTAAGAGACAGTGAAGACCTTCTCCCTGTCATTCAGCTGGC 780
 DB 742 GGATTGAGAACACTGCTAAGAGACAGTGAAGACCTTCTCCCTGTCATTCAGCTGGC 801
 QY 781 TCTGCCATGAGACTATACAGCAGATTAATTTCCCTCTATTCAAAATTAATTCGCAATG 840
 DB 802 TCTGCCATGAGACTATACAGCAGATTAATTTCCCTCTATTCAAAATTAATTCGCAATG 861
 QY 841 AAAATTGGGAATGTAGCATA 861
 DB 862 AAAATTGGGAATGTAGCATA 882

RESULT 3
 AR240728 865 bp DNA 1linear PAT 20-DEC-2002
 LOCUS AR240728
 DEFINITION Sequence 1 from patent US 6468776.
 ACCESSION AR240728
 VERSION AR240728.1 GI:27285916
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 865)
 AUTHORS Conklin,D.C.
 TITLE Human serine protease
 JOURNAL Patent: US 6468776-A 1 22-OCT-2002;
 FEATURES
 source 1..865
 /organism="unknown"

BASE COUNT 235 a 226 c 196 g 208 t
 ORIGIN
 Query Match 98.2%; Score 850.8; DB 6; Length 865;
 Best Local Similarity 99.7%; Prid. No. 6,5e-229;
 Matches 863; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 GGCACTTACTCCCTGAGTAAGGGGAGAGAGCTGATCAACCAATGAAATATGTTCTATT 60
 DB 1 GGCACTTACTCCCTGAGTAAGGGGAGAGAGCTGATCAACCAATGAAATATGTTCTATT 60
 QY 61 TGGGTGTCTCGCTGGAGCAATTTTCTTCTGCTACTCATCTGTTGAGAAAGAACCTTG 120
 DB 61 TGGGTGTCTCGCTGGAGCAATTTTCTTCTGCTACTCATCTGTTGAGAAAGAACCTTG 120
 QY 121 CTCCTCAATTTGGTGTACTCAAGTCTCACTTCAACCCCTGTGTGGCGTCTCATCAAC 180
 DB 121 CTCCTCAATTTGGTGTACTCAAGTCTCACTTCAACCCCTGTGTGGCGTCTCATCAAC 180
 QY 181 CCAGCTGGGAGTGGCCAGCTCAGTGCATTTACCAATTCGAAAGTATGCTGGAA 240
 DB 181 CCAGCTGGGAGTGGCCAGCTCAGTGCATTTACCAATTCGAAAGTATGCTGGAA 240
 QY 241 ATTTCAGAGCAGAGTCAAGAGCGGTATGAAACAATTAACCCCATTCAGATGCTCC 300
 DB 241 ATTTCAGAGCAGAGTCAAGAGCGGTATGAAACAATTAACCCCATTCAGATGCTCC 300

QY 301 GCTACTGAACTACAGTCAATAGGCGCCCAAGAGTACCTCATGCTCATCAAGCTGGCTA 360
 DB 301 GCTACTGAACTACAGTCAATAGGCGCCCAAGAGTACCTCATGCTCATCAAGCTGGCTA 360
 QY 361 AGCTTCGATGCTCAATCCCAAGATGTCAGAGCCCTTCCCTGCGCACACCAATGTGAGG 420
 DB 361 AGCTTCGATGCTCAATCCCAAGATGTCAGAGCCCTTCCCTGCGCACACCAATGTGAGG 420
 QY 421 CAGGCACTGTCTGTCTACTCTCAAGTTTGGACTGAGGCCAAGAAACAGTGGCCGACCC 480
 DB 421 CAGGCACTGTCTGTCTACTCTCAAGTTTGGACTGAGGCCAAGAAACAGTGGCCGACCC 480
 QY 481 CTGACTTGGGCGAGAACCTGAGAGGCCCCGCTGATGTCGAGATGCGCAAAAAAG 540
 DB 481 CTGACTTGGGCGAGAACCTGAGAGGCCCCGCTGATGTCGAGATGCGCAAAAAAG 540
 QY 541 AACAGAGAAAAAGCCACAGAGAAATTCCTTATGTTGTAATTTGTAAGTAATTCAGCCGA 600
 DB 541 AACAGAGAAAAAGCCACAGAGAAATTCCTTATGTTGTAATTTGTAAGTAATTCAGCCGA 600
 QY 601 TTTTGGGAGAGGTGGCCGTTGCTACTGTCATCTGCAAGACAGCTCCAGGGAATGAGG 660
 DB 601 TTTTGGGAGAGGTGGCCGTTGCTACTGTCATCTGCAAGACAGCTCCAGGGAATGAGG 660
 QY 661 TGGGCGACTTCATGAGAGGGGAGCTGCGGACTTACACCAATGTTTCAAAATATGATTCCT 720
 DB 661 TGGGCGACTTCATGAGAGGGGAGCTGCGGACTTACACCAATGTTTCAAAATATGATTCCT 720
 QY 721 GGATTGAGAACACTGCTAAGAGACAGTGAAGACCTTCTCCCTGTCATTCAGCTGGC 780
 DB 721 GGATTGAGAACACTGCTAAGAGACAGTGAAGACCTTCTCCCTGTCATTCAGCTGGC 780
 QY 781 TCTGCCATGAGACTATACAGCAGATTAATTTCCCTCTATTCAAAATTAATTCGCAATG 840
 DB 781 TCTGCCATGAGACTATACAGCAGATTAATTTCCCTCTATTCAAAATTAATTCGCAATG 840
 QY 841 AAAATTGGGAATGTAGCATA 866
 DB 841 AAAATTGGGAATGTAGCATA 886

RESULT 4
 AXJ18279 865 bp DNA 1linear PAT 15-DEC-2001
 LOCUS AXJ18279
 DEFINITION Sequence 1 from Patent WO0166771.
 ACCESSION AXJ18279
 VERSION AXJ18279.1 GI:17900929
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE 1
 AUTHORS Conklin,D.C.
 TITLE Human serine protease
 JOURNAL Patent: WO 0166771-A 1 13-SEP-2001;
 ZymoGenetics, Inc. (US)
 FEATURES
 source 1..865
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 42..749
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD19358.1"
 /db_xref="GI:17900930"
 /translation="MKIVFYLGVLACTPPADSVQKEDPAPLYLYLKSHPKCVL
 IKPSVLAIPHICYLPNLKMLGNFGRVADGTEQINPIQIVRWYNSHSAPODDLMT
 IKLAPAMINPVPQPLTLATNVRPGTALSGDLMWSGNSGRHDLRLNLEAPVSD
 RECKTBOGKSHRSLCVFVFSRIREVAIVATVICHDKLOGIEVGHFMGSDVGIY
 TNVYKYVSWIENTARDK"

CDS
 1..865
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 42..749
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD19358.1"
 /db_xref="GI:17900930"
 /translation="MKIVFYLGVLACTPPADSVQKEDPAPLYLYLKSHPKCVL
 IKPSVLAIPHICYLPNLKMLGNFGRVADGTEQINPIQIVRWYNSHSAPODDLMT
 IKLAPAMINPVPQPLTLATNVRPGTALSGDLMWSGNSGRHDLRLNLEAPVSD
 RECKTBOGKSHRSLCVFVFSRIREVAIVATVICHDKLOGIEVGHFMGSDVGIY
 TNVYKYVSWIENTARDK"

BASE COUNT 235 a 226 c 196 g 208 t
ORIGIN
Query Match 98.2%; Score 850.8; DB 6; Length 865;
Best Local Similarity 99.7%; Pred. No. 6.5e-229;
Matches 863; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 1 GGCACTTACTCCTGAGCTAAGGGGGAAGAGCTGATCACCATGAATATGTCTTCTATT 60
DB 1 GGCACTTACTCCTGAGCTAAGGGGGAAGAGCTGATCACCATGAATATGTCTTCTATT 60
QY 61 TGGGTGTCTCGCTGGGACATTTTCTTGTCTGACTCATCTGTTGAGAAAGAACCTTG 120
DB 61 TGGGTGTCTCGCTGGGACATTTTCTTGTCTGACTCATCTGTTGAGAAAGAACCTTG 120
QY 121 CTCCTATTGTTGGTGAACCTCAAGCTCATTCAACCCCTGTGTGGGCTCTCATCAAC 180
DB 121 CTCCTATTGTTGGTGAACCTCAAGCTCATTCAACCCCTGTGTGGGCTCTCATCAAC 180
QY 181 CCAGCTGGGTGTGGGCTGAGCTCACTGCTATTATTCGAATCTGAAGATGATGCTGGAA 240
DB 181 CCAGCTGGGTGTGGGCTGAGCTCACTGCTATTATTCGAATCTGAAGATGATGCTGGAA 240
QY 241 ATTTCAAGAGCAGAGTCAAGACGGTACTGAACAGACAAATTAACCCATTCAAGTCTCC 300
DB 241 ATTTCAAGAGCAGAGTCAAGACGGTACTGAACAGACAAATTAACCCATTCAAGTCTCC 300
QY 301 GCTACTGGAACTCAAGTCAAGTCAAGGCCCCCAGAGATGATCTCATCAAGCTGCTA 360
DB 301 GCTACTGGAACTCAAGTCAAGTCAAGGCCCCCAGAGATGATCTCATCAAGCTGCTA 360
QY 361 AGCTTCGATGTCTCAATCCCAAGTCCAGCCCTTCCCTGCGCACCAACCAATGTCAAGC 420
DB 361 AGCTTCGATGTCTCAATCCCAAGTCCAGCCCTTCCCTGCGCACCAACCAATGTCAAGC 420
QY 421 CAGGCACTGTCTGTCTACTCTCAAGTTTGAAGCTGAGCCAAAGAAACAGTGGCCGAC 480
DB 421 CAGGCACTGTCTGTCTACTCTCAAGTTTGAAGCTGAGCCAAAGAAACAGTGGCCGAC 480
QY 481 CTGACTTGGCGGAGAACTTGGAGGCCCCCGTGTGATGATGAGAAATGCAAAAAACAG 540
DB 481 CTGACTTGGCGGAGAACTTGGAGGCCCCCGTGTGATGATGAGAAATGCAAAAAACAG 540
QY 541 AACAGAAAAAGCCACAGGAATTCCTTATGTGTGAATTTGTGAAGATTTGAGCCGAA 600
DB 541 AACAGAAAAAGCCACAGGAATTCCTTATGTGTGAATTTGTGAAGATTTGAGCCGAA 600
QY 601 TTTTGGGGAGGTGGCCGTTGCTACTGTCTGCAAGAACAGCTCCAGGGAATCGAG 660
DB 601 TTTTGGGGAGGTGGCCGTTGCTACTGTCTGCAAGAACAGCTCCAGGGAATCGAG 660
QY 661 TGGGGCACTTCAATGGAGGGGAGCTGGGCTTACACCAATGTTTACCAATATGTATCT 720
DB 661 TGGGGCACTTCAATGGAGGGGAGCTGGGCTTACACCAATGTTTACCAATATGTATCT 720
QY 721 GGAATGAAACATCTGCTAAGGACAAGTGAACCTTCTCCCTGCAATTCGATCGC 780
DB 721 GGAATGAAACATCTGCTAAGGACAAGTGAACCTTCTCCCTGCAATTCGATCGC 780
QY 781 TCTGCAATGACTATACAGCAGATATTTTCCCTTATTCAAATTAATTCCTCAATG 840
DB 781 TCTGCAATGACTATACAGCAGATATTTTCCCTTATTCAAATTAATTCCTCAATG 840
QY 841 AAAATTTGGGAATGTAGCAATCTAGT 866
DB 841 AAAATTTGGGAATGTAGCAATCTAGT 866

RESULT 5
AR263885/c AR263885 865 bp DNA linear PAT 29-JAN-2003
LOCUS AR263885
DEFINITION Sequence 63 from patent US 6331427.
ACCESSION AR263885

VERSION AR263885.1 GI:28075889
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 865)
TITLES Robison, K.E.
JOURNAL Protease homologs
FEATURES Patent: US 6331427-A 63 18-DEC-2001;
Location/Qualifiers
source 1..865
BASE COUNT 207 a 197 c 225 g 236 t
ORIGIN
Query Match 96.3%; Score 834; DB 6; Length 865;
Best Local Similarity 98.7%; Pred. No. 3.6e-224;
Matches 851; Conservative 0; Mismatches 10; Indels 1; Gaps 1;
QY 1 GGCACTTACTCCTGAGCTAAGGGGGAAGAGCTGATCACCATGAATATGTCTTCTATT 60
DB 865 GGCACTTACTCCTGAGCTAAGGGGGAAGAGCTGATCACCATGAATATGTCTTCTATT 866
QY 61 TGGGTGTCTCGCTGGGACATTTTCTTGTCTGACTCATCTGTTGAGAAAGAACCTTG 120
DB 805 TGGGTGTCTCGCTGGGACATTTTCTTGTCTGACTCATCTGTTGAGAAAGAACCTTG 746
QY 121 CTCCTATTGTTGGTGAACCTCAAGCTCATTCAACCCCTGTGTGGGCTCTCATCAAC 180
DB 745 CTCCTATTGTTGGTGAACCTCAAGCTCATTCAACCCCTGTGTGGGCTCTCATCAAC 180
QY 181 CCAGCTGGGTGTGGGCTGAGCTCACTGCTATTATTCGAATCTGAAGATGATGCTGGAA 240
DB 685 CCAGCTGGGTGTGGGCTGAGCTCACTGCTATTATTCGAATCTGAAGATGATGCTGGAA 240
QY 241 ATTTCAAGAGCAGAGTCAAGACGGTACTGAACAGACAAATTAACCCATTCAAGTCTCC 300
DB 625 ATTTCAAGAGCAGAGTCAAGACGGTACTGAACAGACAAATTAACCCATTCAAGTCTCC 300
QY 301 GCTACTGGAACTCAAGTCAAGTCAAGGCCCCCAGAGATGATCTCATCAAGCTGCTA 360
DB 565 GCTACTGGAACTCAAGTCAAGTCAAGGCCCCCAGAGATGATCTCATCAAGCTGCTA 360
QY 361 AGCTTCGATGTCTCAATCCCAAGTCCAGCCCTTCCCTGCGCACCAACCAATGTCAAGC 419
DB 505 AGCTTCGATGTCTCAATCCCAAGTCCAGCCCTTCCCTGCGCACCAACCAATGTCAAGC 419
QY 420 CAGGCACTGTCTGTCTACTCTCAAGTTTGAAGCTGAGCCAAAGAAACAGTGGCCGAC 479
DB 445 CAGGCACTGTCTGTCTACTCTCAAGTTTGAAGCTGAGCCAAAGAAACAGTGGCCGAC 479
QY 480 CCGTACTTGGCGGAGAACTTGGAGGCCCCCGTGTGATGATGAGAAATGCAAAAAACAG 539
DB 385 CCGTACTTGGCGGAGAACTTGGAGGCCCCCGTGTGATGATGAGAAATGCAAAAAACAG 539
QY 540 GAAACAAGAAAAAGCCACAGGAATTCCTTATGTGTGAATTTGTGAAGATTTGAGCCGAA 599
DB 325 GAAACAAGAAAAAGCCACAGGAATTCCTTATGTGTGAATTTGTGAAGATTTGAGCCGAA 599
QY 600 ATTTTGGGGAGGTGGCCGTTGCTACTGTCTGCAAGAACAGCTCCAGGGAATCGAG 659
DB 265 ATTTTGGGGAGGTGGCCGTTGCTACTGTCTGCAAGAACAGCTCCAGGGAATCGAG 659
QY 660 GTGGGGCACTTCAATGGAGGGGAGCTGGGCTTACACCAATGTTTACCAATATGTATCT 719
DB 205 GTGGGGCACTTCAATGGAGGGGAGCTGGGCTTACACCAATGTTTACCAATATGTATCT 719
QY 720 TGGATGAGAAACATCTGCTAAGGACAAGTGAACCTTCTCCCTGCAATTCGATCGC 779
DB 145 TGGATGAGAAACATCTGCTAAGGACAAGTGAACCTTCTCCCTGCAATTCGATCGC 779
QY 780 CTCCTGCAATGACTATACAGCAGATATTTTCCCTTATTCAAATTAATTCCTCAATG 839
DB 780 CTCCTGCAATGACTATACAGCAGATATTTTCCCTTATTCAAATTAATTCCTCAATG 839

Db 85 CTCCTGCATGAGATATACAGACAGATATTTCTCTATTCATAATTAATTCCTCAAT 26
QY 840 GAAATTTGGGAATGTAGCAT 861
Db 25 GAAATTTGGGAATGTAGCAA 4

RESULT 6
AX350386
LOCUS AX350386 730 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 11 from Patent WO0181578.
ACCESSION AX350386
VERSION AX350386.1 GI:18616037
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Verne, C.A., Fernandes, E.R., Gerlach, V., Shinkets, R.A.,
Malpanker, U.M., Boldog, F.L., Zerkusen, B.D., Spyrek, K.A.,
Majumder, K., Tchernev, V.T., Padigaru, M., Patutajan, M.,
Burgess, C.E., Gangoli, E.A., Smithson, G., Rastelli, L.,
Maddougall, J.R., Taupier, R.J., Grose, W.M. and Alsobrook, J.P.
TITLE Novel proteins and nucleic acids encoding same
JOURNAL Patent: WO 0181578-A 11 01-NOV-2001;
Curegen Corporation (US)
FEATURES
source Location/Qualifiers
1..730
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 191 a 197 c 170 g 172 t
ORIGIN

Query Match 84.3%; Score 730; DB 6; Length 730;
Best Local Similarity 100.0%; Pred. No. 8.5e-195;
Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 GATCACCATGAATATATGTCTTATTTGGGTCTCTGCTGGACATTTTCTTCTGA 94
Db 1 GATCACCATGAATATATGTCTTATTTGGGTCTCTGCTGGACATTTTCTTCTGA 60
QY 95 CTCATCTGTTCAAGAAAGAACCCCTGCTCTATTTGGGTCTCTGCTGGACATTTTCTTCTGA 154
Db 61 CTCATCTGTTCAAGAAAGAACCCCTGCTCTATTTGGGTCTCTGCTGGACATTTTCTTCTGA 120
QY 155 CCCCTGTGTGGGCGTCTCTATCAACCCAGCTGGGTGGCCCACTCATGCTATTT 214
Db 121 CCCCTGTGTGGGCGTCTCTATCAACCCAGCTGGGTGGCCCACTCATGCTATTT 180
QY 215 ACCAATCTGAAGATGCTGGAAATTTCAAGACAGACAGTCAAGACGCTACTGA 274
Db 181 ACCAATCTGAAGATGCTGGAAATTTCAAGACAGACAGTCAAGACGCTACTGA 240
QY 275 GACAAATTAACCCATTCAGATGCTCGCTAAGTCAAGTCAAGGCCCCCAAGGA 334
Db 241 GACAAATTAACCCATTCAGATGCTCGCTAAGTCAAGTCAAGGCCCCCAAGGA 300
QY 335 TGACCTCATGCTCATCAAGCTGCTCAAGCTGCTCAATCCCAAGTCCAGCCCT 394
Db 301 TGACCTCATGCTCATCAAGCTGCTCAAGCTGCTCAATCCCAAGTCCAGCCCT 360
QY 395 TCCCTCTGCGCACCAATATCTCAAGGCGACCTGTCTGTCTAATCTCAAGTTTGA 454
Db 361 TCCCTCTGCGCACCAATATCTCAAGGCGACCTGTCTGTCTAATCTCAAGTTTGA 420
QY 455 GAGCCAAAGAAAGTGGCGAGACCCCTGACTGGCGAGACCTGAGAGGCCCCCGTGA 514
Db 421 GAGCCAAAGAAAGTGGCGAGACCCCTGACTGGCGAGACCTGAGAGGCCCCCGTGA 480
QY 515 GTCTGATGAGATGCGCAAAAAACAGAACAGAAAAAGCAGACAGAAATTCCTATGTGT 574

Db 481 GTCTGATCGAAGATGCGCAAAAAACAGAACAGAAAAAGCCACAGAAATTCCTATGTGT 540
QY 575 GAAATTTGGAAGATATTCAGCGCAATTTTGGGGAGGTGGCGGTGCTACTGTCTG 634
Db 541 GAAATTTGGAAGATATTCAGCGCAATTTTGGGGAGGTGGCGGTGCTACTGTCTG 600
QY 635 CAAAGCAAGCTCCAGGAATGAGGTGGGCACTTCATGAGAGGAGCGTGGCATCTA 694
Db 601 CAAAGCAAGCTCCAGGAATGAGGTGGGCACTTCATGAGAGGAGCGTGGCATCTA 660
QY 695 CACCAATGTTTCAATATATGATCTCTGATTTGAGAACATCTGTAAGACAAAGTGA 754
Db 661 CACCAATGTTTCAATATATGATCTCTGATTTGAGAACATCTGTAAGACAAAGTGA 720
QY 755 TACTTCTGCC 764
Db 721 TACTTCTGCC 730

RESULT 7
AX350388
LOCUS AX350388 730 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 13 from Patent WO0181578.
ACCESSION AX350388
VERSION AX350388.1 GI:18616038
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Verne, C.A., Fernandes, E.R., Gerlach, V., Shinkets, R.A.,
Malpanker, U.M., Boldog, F.L., Zerkusen, B.D., Spyrek, K.A.,
Majumder, K., Tchernev, V.T., Padigaru, M., Patutajan, M.,
Burgess, C.E., Gangoli, E.A., Smithson, G., Rastelli, L.,
Maddougall, J.R., Taupier, R.J., Grose, W.M. and Alsobrook, J.P.
TITLE Novel proteins and nucleic acids encoding same
JOURNAL Patent: WO 0181578-A 13 01-NOV-2001;
Curegen Corporation (US)
FEATURES
source Location/Qualifiers
1..730
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 191 a 198 c 170 g 171 t
ORIGIN

Query Match 84.1%; Score 728.4; DB 6; Length 730;
Best Local Similarity 99.9%; Pred. No. 2.4e-194;
Matches 729; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 35 GATCACCATGAATATATGTCTTATTTGGGTCTCTGCTGGACATTTTCTTCTGA 94
Db 1 GATCACCATGAATATATGTCTTATTTGGGTCTCTGCTGGACATTTTCTTCTGA 60
QY 95 CTCATCTGTTCAAGAAAGAACCCCTGCTCTATTTGGGTCTCTGCTGGACATTTTCTTCTGA 154
Db 61 CTCATCTGTTCAAGAAAGAACCCCTGCTCTATTTGGGTCTCTGCTGGACATTTTCTTCTGA 120
QY 155 CCCCTGTGTGGGCGTCTCTATCAACCCAGCTGGGTGGCCCACTCATGCTATTT 214
Db 121 CCCCTGTGTGGGCGTCTCTATCAACCCAGCTGGGTGGCCCACTCATGCTATTT 180
QY 215 ACCAATCTGAAGATGCTGGAAATTTCAAGACAGACAGTCAAGACGCTACTGA 274
Db 181 ACCAATCTGAAGATGCTGGAAATTTCAAGACAGACAGTCAAGACGCTACTGA 240
QY 275 GACAAATTAACCCATTCAGATGCTCGCTAAGTCAAGTCAAGGCCCCCAAGGA 334
Db 241 GACAAATTAACCCATTCAGATGCTCGCTAAGTCAAGTCAAGGCCCCCAAGGA 300
QY 335 TGACCTCATGCTCATCAAGCTGCTCAAGCTGCTCAATCCCAAGTCCAGCCCT 394

Db 301 TGACCTATGCTATCAAGCTGGCTAAGCTCCATCTCAATCCAAAGTCAGCCCT 360
QY 395 TCCCTCGCCACCAACCAATGTCAGGCGACCTGTCTCTCACTCAGATTGACTG 454
Db 361 TCCCTCGCCACCAACCAATGTCAGGCGACCTGTCTCTCACTCAGATTGACTG 420
QY 455 GAGCCAGAAAACAGTGGCCGACACCTGACTTGGCGAGAACCTGGAGGCCCGTGT 514
Db 421 GAGCCAGAAAACAGTGGCCGACACCTGACTTGGCGAGAACCTGGAGGCCCGTGT 480
QY 515 GTCTGATCGAAGATGCCAAAACAGAAACAGAAAGCCAGCAAGATTCCTTAATGT 574
Db 481 GTCTGATCGAAGATGCCAAAACAGAAACAGAAAGCCAGCAAGATTCCTTAATGT 540
QY 575 GAAATTTGTAAAGATATTCAGCCGAATTTTGGGAGGTGGCCCTTCTACTGCTG 634
Db 541 GAAATTTGTAAAGATATTCAGCCGAATTTTGGGAGGTGGCCCTTCTACTGCTG 600
QY 635 CAAAGACACTCCAGGAAATCGAGTGGGCACTTCATGAGAGGGAAGCTGGCATCTA 694
Db 601 CAAAGACACTCCAGGAAATCGAGTGGGCACTTCATGAGAGGGAAGCTGGCATCTA 660
QY 695 CACCAATGTTTAAATATATCTGATTTGAGAACACTGCTAAGACAAAGTGAAGCC 754
Db 661 CACCAATGTTTAAATATATCTGATTTGAGAACACTGCTAAGACAAAGTGAAGCC 720
QY 755 TACTCTCC 764
Db 721 TACTCTCC 730

RESULT 8
AX247864 708 bp DNA linear PAT 28-SEP-2001
LOCUS Sequence 105 from Patent WO0166748.
DEFINITION AX247864
ACCESSION AX247864 GI:15862497
VERSION AX247864.1 GI:15862497
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Conklin,D.C., Presnell,S.R. and Adler,D.A.
TITLE Full length expressed human polynucleotides and the polypeptides
they encode
JOURNAL Patent: WO 0166748-A,105 13-SEP-2001;
ZymoGenetics, Inc. (US)

FEATURES
source Location/Qualifiers
1..708
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

CDS
1..708
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC8648.1"
/db_xref="GI:15862498"

FEATURES
source Location/Qualifiers
1..708
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 188 a 185 c 168 g 167 t
ORIGIN
Query Match 81.6%; Score 706.4; DB 6; Length 708;
Best Local Similarity 99.9%; Pred. No. 3.9e-188;
Matches 707; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 42 ATGAATATGCTCTTATTTGGGTGCTCGTGGGACATTTTCTTGTGACTCATCT 101
Db 1 ATGAATATGCTCTTATTTGGGTGCTCGTGGGACATTTTCTTGTGACTCATCT 60

QY 102 GTTCAGAAAGAACCCCTGCTCCCTATTGTTGTTACTTCAAGTCTCACTTCAACCCCTGT 161
Db 61 GTTCAGAAAGAACCCCTGCTCCCTATTGTTGTTACTTCAAGTCTCACTTCAACCCCTGT 120
QY 162 GTGGGGTCTCTCATCAAAACCCAGCTGGTGTGTCGGCCCAAGCTACTGCTATTTACCAAT 221
Db 121 GTGGGGTCTCTCATCAAAACCCAGCTGGTGTGTCGGCCCAAGCTACTGCTATTTACCAAT 180
QY 222 CTGAAGATGATGTGGAAATTTCAAGAGAGAGTCAAGAGAGGTCTGAAGACCAATT 281
Db 181 CTGAAGATGATGTGGAAATTTCAAGAGAGAGTCAAGAGAGGTCTGAAGACCAATT 240
QY 282 AACCCATTCAGATCGTCCGCTACTGAACTACATAGCCGCCACAGATGACCTC 341
Db 241 AACCCATTCAGATCGTCCGCTACTGAACTACATAGCCGCCACAGATGACCTC 300
QY 342 ATGCTCATCAAGTGGCTTAAGCTGCAATGCTCAATCCAAAGTCCAGCCCTTCCCTC 401
Db 301 ATGCTCATCAAGTGGCTTAAGCTGCAATGCTCAATCCAAAGTCCAGCCCTTCCCTC 360
QY 402 GCCACCAACCAATGTCAGGCGAGGCACTGTCTGTCTCACTGAGGTTGACTGAGCCAA 461
Db 361 GCCACCAACCAATGTCAGGCGAGGCACTGTCTGTCTCACTGAGGTTGACTGAGCCAA 420
QY 462 GAAACAGTGGCCGACACCTGACTTGGGCAAGAACCTGAGGCCCCGCTGATGCTGAT 521
Db 421 GAAACAGTGGCCGACACCTGACTTGGGCAAGAACCTGAGGCCCCGCTGATGCTGAT 480
QY 522 CGAAGATGCCAAAACAGAACAGAAAGCCACAGAAATTCCTTATGTGAAATTT 581
Db 481 CGAAGATGCCAAAACAGAACAGAAAGCCACAGAAATTCCTTATGTGAAATTT 540
QY 582 GTGAAGATTTACGCGGAATTTTGGGAGGTGGGCGGTGCTACTGCTCAATGCAAGAC 641
Db 541 GTGAAGATTTACGCGGAATTTTGGGAGGTGGGCGGTGCTACTGCTCAATGCAAGAC 600
QY 642 AAGCTCAGGGAATCGAGGTGGGCACTTCATGAGAGGAGGAGCGTCCGCACTTACACCAAT 701
Db 601 AAGCTCAGGGAATCGAGGTGGGCACTTCATGAGAGGAGGAGCGTCCGCACTTACACCAAT 660
QY 702 GTTTACAAATATGATATCTGATTTGAGAAACACTGCTTAAGAGCAAGTGA 749
Db 661 GTTTACAAATATGATATCTGATTTGAGAAACACTGCTTAAGAGCAAGTGA 708

RESULT 9
AX455326 708 bp DNA linear PAT 06-JUL-2002
LOCUS Sequence 3 from Patent WO0206331.
DEFINITION AX455326
ACCESSION AX455326
VERSION AX455326.1 GI:21714509
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Meyers,R.
TITLE 14087, a novel serine protease molecule and uses therefor
JOURNAL Patent: WO 0206331-A 3 24-JAN-2002;
Millennium Pharmaceuticals, Inc. (US)

FEATURES
source Location/Qualifiers
1..708
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 186 a 187 c 168 g 167 t
ORIGIN
Query Match 81.6%; Score 706.4; DB 6; Length 708;
Best Local Similarity 99.9%; Pred. No. 3.9e-188;
Matches 707; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 42 ATGAATATGCTCTTCTATTGAGTGTCTCGCTGGACATTTTCTTCTGCTCATCT 101
DB 1 ATGAATATGCTCTTCTATTGAGTGTCTCGCTGGACATTTTCTTCTGCTCATCT 60
QY 102 GTTCAGAAAGAACCCCTGCTCTTCTTGTGTGATCTCAAGTCTCATCTCAACCCCTGT 161
DB 61 GTTCAGAAAGAACCCCTGCTCTTCTTGTGTGATCTCAAGTCTCATCTCAACCCCTGT 120
QY 162 GTGGGGGCTCTCATCAACCCAGCTGGGTGCTGGCCCGCTCACTGCTATTACCAAT 221
DB 121 GTGGGGGCTCTCATCAACCCAGCTGGGTGCTGGCCCGCTCACTGCTATTACCAAT 180
QY 222 CTGAAAGTATCTGCGGAATTTTCAAGACAGAGTCAAGACGCTACTGAACAGCAAT 281
DB 181 CTGAAAGTATCTGCGGAATTTTCAAGACAGAGTCAAGACGCTACTGAACAGCAAT 240
QY 282 AACCCCATTCAGATGCTCGCTACTGGAACCTACAGTATAGCGCCACAGATGACCTC 341
DB 241 AACCCCATTCAGATGCTCGCTACTGGAACCTACAGTATAGCGCCACAGATGACCTC 300
QY 342 ATGCTCATCAAGCTGCTAAGCTGCTGCACTCAATCCCAAGTCCAGCCCTTCCCTTC 401
DB 301 ATGCTCATCAAGCTGCTAAGCTGCTGCACTCAATCCCAAGTCCAGCCCTTCCCTTC 360
QY 402 GCCACCAACCAATGTCAGGCGCAGCACTGTCTGTCTACTCTCAAGTTTGAAGTGAAGCA 461
DB 361 GCCACCAACCAATGTCAGGCGCAGCACTGTCTGTCTACTCTCAAGTTTGAAGTGAAGCA 420
QY 462 GAAAAAGTGGCGCAGCACTGTGCTGCGGAGAACCTGGAAGGCCCCGCTGATGTGAT 521
DB 421 GAAAAAGTGGCGCAGCACTGTGCTGCGGAGAACCTGGAAGGCCCCGCTGATGTGAT 480
QY 522 CGAGAAATGCGCAAAAAACAAGAAAAAGCAAGAAATTCCTTAATGTGAAATTT 581
DB 481 CGAGAAATGCGCAAAAAACAAGAAAAAGCAAGAAATTCCTTAATGTGAAATTT 540
QY 582 GTGAAGATTTTCAGCGGAATTTTGGGAGGTGGCCGTTGCTACTGCTCACTGCAAGAAC 641
DB 541 GTGAAGATTTTCAGCGGAATTTTGGGAGGTGGCCGTTGCTACTGCTCACTGCAAGAAC 600
QY 642 AAGCTTCAGGGAATCGAGGTGGGCACTTCATGGAAGGGAAGTCCGCACTCAACCAAT 701
DB 601 AAGCTTCAGGGAATCGAGGTGGGCACTTCATGGAAGGGAAGTCCGCACTCAACCAAT 660
QY 702 GTTTCACAAATATGATCTCTGATGGAACACTGCTAAGACAGATGA 749
DB 661 GTTTCACAAATATGATCTCTGATGGAACACTGCTAAGACAGATGA 708

RESULT 10
AX350481 721 bp DNA linear PART 06-FEB-2002
LOCUS AX350481
DEFINITION Sequence 106 from Patent WO0181578.
ACCESSION AX350481
VERSION AX350481.1 GI:18616087
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 Vernet,C.A., Fernandes,E.R., Gerlach,V., Shinkets,R.A.,
Malyanar,U.M., Boldog,F.L., Zernhusen,B.D., Spytek,K.A.,
Majumder,K., Tchertnev,V.T., Padigara,M., Patuzaj,M.,
Burgess,C.E., Gangoli,E.A., Smlthson,G., Rastelli,L.,
Macdonnell,J.R., Taupier,R.J., Grose,W.M. and Alsobrook,J.P.
Novel proteins and nucleic acids encoding same
Patent: WO 0181578-A 106 01-NOV-2001;
LOCATION/Qualifiers
1..721
FEATURES
Source
/organism="Homo sapiens"

```

```

BASE COUNT 193 a 191 c 169 g 168 t
ORIGIN
Query Match 81.3%; Score 703.8; DB 6; Length 721;
Best Local Similarity 99.6%; Pred. No. 2.1e-187;
Matches 716; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 39 ACCATGAATATGCTCTTCTATTGAGTGTCTCGCTGGACATTTTCTTCTGCTCATCT 98
DB 2 ACCATGAATATGCTCTTCTATTGAGTGTCTCGCTGGACATTTTCTTCTGCTCATCT 61
QY 99 TCTGTTCAGAAAGAACCCCTGCTCTTCTTGTGTGATCTCAAGTCTCATCTCAACCC 158
DB 62 TCTGTTCAGAAAGAACCCCTGCTCTTCTTGTGTGATCTCAAGTCTCATCTCAACCC 121
QY 159 TGTGTGGGGCTCTCATCAACCCAGCTGGGTGCTGGCCCGCTCACTGCTATTACCA 218
DB 122 TGTGTGGGGCTCTCATCAACCCAGCTGGGTGCTGGCCCGCTCACTGCTATTACCA 181
QY 219 AATCTGAAGTATGCTGCGGAATTTTCAAGACAGAGTCAAGACGCTACTGAACAGACA 278
DB 182 AATCTGAAGTATGCTGCGGAATTTTCAAGACAGAGTCAAGACGCTACTGAACAGACA 241
QY 279 ATTAAACCCATTCAGATGCTCGCTACTGGAACCTACAGTATAGGCGCCACAGATGAC 338
DB 242 ATTAAACCCATTCAGATGCTCGCTACTGGAACCTACAGTATAGGCGCCACAGATGAC 301
QY 339 CTGATGCTCATCAAGCTGCTAAGCTGCTGCACTCAATCCCAAGTCCAGCCCTTCC 398
DB 302 CTGATGCTCATCAAGCTGCTAAGCTGCTGCACTCAATCCCAAGTCCAGCCCTTCC 361
QY 399 CTGCGCACCAACCAATGTCAGGCGCAGCACTGTCTGTCTACTCTCAAGTTTGAAGTGAAG 458
DB 362 CTGCGCACCAACCAATGTCAGGCGCAGCACTGTCTGTCTACTCTCAAGTTTGAAGTGAAG 421
QY 459 CAAGAAACAGTGGCGCAGCACTGCTGCTGCGGAGAACCTGGAAGGCCCCGCTGATGTCT 518
DB 422 CAAGAAACAGTGGCGCAGCACTGCTGCTGCGGAGAACCTGGAAGGCCCCGCTGATGTCT 481
QY 519 GATCAGAAATGCC-AAAAACAGAAACAAGAAAAAGCCACAGAAATTCCTTAATGTGAA 577
DB 482 GATCAGAAATGCC-AAAAACAGAAACAAGAAAAAGCCACAGAAATTCCTTAATGTGAA 541
QY 578 ATTGTGAAAGTATTCAGCCGAATTTTGGGAGGTGGCCGTTGCTACTGCTCATCTGCAA 637
DB 542 ATTGTGAAAGTATTCAGCCGAATTTTGGGAGGTGGCCGTTGCTACTGCTCATCTGCAA 601
QY 638 AGACAAGCTCCAGGGAATCGAGGTGGGCACTTCATGGAAGGGAAGTCCGCACTCAAC 697
DB 602 AGACAAGCTCCAGGGAATCGAGGTGGGCACTTCATGGAAGGGAAGTCCGCACTCAAC 661
QY 698 CAATGTTTACAAATATGATCTCTGATGGAACACTGCTAAGACAGATGAGACCCCTA 756
DB 662 CAATGTTTACAAATATGATCTCTGATGGAACACTGCTAAGACAGATGAGACCCCTA 720

RESULT 11
AX360087 798 bp DNA linear PART 13-FEB-2002
LOCUS AX360087
DEFINITION Sequence 43 from Patent WO0200860.
ACCESSION AX360087
VERSION AX360087.1 GI:18675713
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 Plozman,G., Whyte,D., Sudarsanam,S., Manning,G., Caenepeel,S. and
Charzydzak,G.
Novel proteases
TITLE

```

JOURNAL Patent: WO 0200860-A 43 03-JAN-2002;
Sugen, Inc. (US)
Location/Qualifiers
source 1. 798
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 212 a 212 c 191 g 183 t
ORIGIN

Query Match 70.0%; Score 606.4; DB 6; Length 798;
Best Local Similarity 88.4%; Pred. No. 6.9e-160;
Matches 707; Conservative 0; Mismatches 1; Indels 90; Gaps 1;

42 ATGAATATGCTCTTCTTATTTGGGTCCTCGTGGGACATTTTCTTCTGACTCATCT 101
1 ATGAATATGCTCTTCTTATTTGGGTCCTCGTGGGACATTTTCTTCTGACTCATCT 60
102 GTTCAGAAAGAACCTGCTCCCTATTTGGTGTACTCAAGTCTCACTTCAACCCCTGT 161
61 GTTCAGAAAGAACCTGCTCCCTATTTGGTGTACTCAAGTCTCACTTCAACCCCTGT 120
162 GTGGGCTCTCTCATCAACCCAGTGGGTGCTGGCCCACTGCTCTATTTACCAAT 221
121 GTGGGCTCTCTCATCAACCCAGTGGGTGCTGGCCCACTGCTCTATTTACCAAT 180
222 CTGAAGATGATGCTGGGAATTTCAAGACAGAGTCAAGACGCTACTGAACAGCAAT 281
181 CTGAAGATGATGCTGGGAATTTCAAGACAGAGTCAAGACGCTACTGAACAGCAAT 240
282 AACCCCATTCAGATCGTCGCTACTGAGAACTACAGTATAGCCGCCACAGATGACTC 341
241 AACCCCATTCAGATCGTCGCTACTGAGAACTACAGTATAGCCGCCACAGATGACTC 300
342 ATGCTCATCAAGCTGGTAAAGCTGCAATGCTCAATCCCAAGTCCAGCCCTTCCCTC 401
301 ATGCTCATCAAGCTGGTAAAGCTGCAATGCTCAATCCCAAGTCCAGCCCTTCCCTC 360
402 GCCACCAACCAATGTCAGAGCCAGGCACTGCTGCTACTCTCAGGTTTGAAGTGAAGCA 461
361 GCCACCAACCAATGTCAGAGCCAGGCACTGCTGCTACTCTCAGGTTTGAAGTGAAGCA 420
462 GAAACAGT----- 470
421 GAAACAGTGGGCTTTGGCAGCTGAGCACAAGCCATCTGACAGAGGCCA 480
471 -----GCCGACACCTTGACTTGGG 491
481 GCCATTCCTGATTGGCAGAGACCAATTCACATGAACAAAGCCGACCTTGACTTGGG 540
492 CAGAACCTGAGAGCCCGCTGATGTCATGAGATGCAAGAAACAGAAACAGAAACAGAA 551
541 CAGAACCTGAGAGCCCGCTGATGTCATGAGATGCAAGAAACAGAAACAGAAACAGAA 600
552 AGCCACAGAAATCTTATGTTGAAATTTGTGAAGTATTCAGCCGAATTTTGGGAG 611
601 AGCCACAGAAATCTTATGTTGAAATTTGTGAAGTATTCAGCCGAATTTTGGGAG 660
612 GTGGCCGCTGCTACTGTCATCTGCAAGACCAAGTCCAGGGAATCCAGGTGGGCACTTC 671
661 GTGGCCGCTGCTACTGTCATCTGCAAGACCAAGTCCAGGGAATCCAGGTGGGCACTTC 720
672 ATGGAGAGGGAGCGTCGGCATCTACCAAGTGTTCAGAAATATGATCTGGAATGGAAC 721
721 ATGGAGAGGGAGCGTCGGCATCTACCAAGTGTTCAGAAATATGATCTGGAATGGAAC 780
732 ACTGCTAAGGACAAAGTGA 749
781 ACTGCTAAGGACAAAGTGA 798

RESULT 12
AK097648

LOCUS AK097648 1855 bp mRNA linear PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FLJ40329 fis, clone TEST12031418, weakly similar
to TRPSIN I-P1 PRECURSOR (EC 3.4.21.4).
ACCESSION AK097648
VERSION AK097648.1 GI:21757486
KEYWORDS oligo capping, file (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS 1 Kanehori, K., Ishibashi, T., Chiba, Y., Fujimori, K., Hiraoka, S.,
Tanai, H., Watanabe, S., Ishida, S., Ono, Y., Houta, T., Watanabe, M.,
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
Kimura, K., Maruo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
Wagatsuna, M., Takahashi-Pujili, A., Oshima, A., Sugiyama, A.,
Kawakami, B., Suzuki, Y., Sugano, S., Nishihara, K., Maehno, Y., Nagai, K.
and Isogai, T.
TITLE MEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1855)
AUTHORS Isogai, T. and Yamamoto, J.
TITLE Direct Submission
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
COMMENT (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
MEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology (RAB); cDNA library
construction; Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5' - 3' end one pass sequencing; RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing; HRI and
RAB; annotation: HRI and RAB.
FEATURES
source
1. 1855
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TEST12031418"
/issue_type="testis"
/clone_lib="TEST12"
/note="Cloning vector: pMB18SFL3"
BASE COUNT 564 a 438 c 408 g 445 t
ORIGIN

Query Match 60.9%; Score 527; DB 9; Length 1855;
Best Local Similarity 87.4%; Pred. No. 2e-137;
Matches 627; Conservative 0; Mismatches 0; Indels 90; Gaps 1;

218 AATCTGAAGATGATGCTGGGAAATTTCAAGACAGAGTCAAGACGCTAAGACGAC 277
1123 AATCTGAAGATGATGCTGGGAAATTTCAAGACAGAGTCAAGACGCTAAGACGAC 1182
278 AATTAACCCCATTCAGATGCTCCGCTACTGAAGTCAAGTCAAGAGCCCAAGATGA 337
1183 AATTAACCCCATTCAGATGCTCCGCTACTGAAGTCAAGTCAAGAGCCCAAGATGA 1242
338 CCTCATGCTCATCAAGCTGGCTAAGCTGCGCATGCTCAATCCAAAGTCCAGGCCCTTCC 397
1243 CCTCATGCTCATCAAGCTGGCTAAGCTGCGCATGCTCAATCCAAAGTCCAGGCCCTTCC 1302
398 CCTGCGCACCAACCAAGTCAAGGCAAGTCTGCTCACTTCAAGTTTGAAGTGAAG 457
1303 CCTGCGCACCAACCAAGTCAAGGCAAGTCTGCTCACTTCAAGTTTGAAGTGAAG 1362
458 CCAAGAAAACAGT----- 470
1363 CCAAGAAAACAGTGGGCTTTGGCAGCTGAGACCAAGCCATCTGACAGAGG 1422
471 -----GCCGACACCTTGACTT 487

Db 1423 CCAGCCATTCTGATTTGGAGAGACAAATTCATGAACAAAGCCGACCCCTGACTT 1482
Qy 488 GCGGCAAGACCTTGAAGGCCCCCGGATGTCGATTCGAATGCAAAAACAGAACAGG 547
Db 1483 GCGGCAAGACCTTGAAGGCCCCCGGATGTCGATTCGAATGCAAAAACAGAACAGG 1542
Qy 548 AAAAGCCACAGAAATTCCTATGTCGAAATTTGTGAAATTTGCAAGCCGAATTTTGG 607
Db 1543 AAAAGCCACAGAAATTCCTATGTCGAAATTTGTGAAATTTGCAAGCCGAATTTTGG 1602
Qy 608 GAGGTCGCGCTTCTACTGTCATCTGCAAAAGACAGCTCCAGGAAATGAGTGGGCA 667
Db 1603 GAGGTCGCGCTTCTACTGTCATCTGCAAAAGACAGCTCCAGGAAATGAGTGGGCA 1662
Qy 668 CTTCATGAGGAGGAGGAGTGGGATCTACACCAATGTTTAAATATGATCTGATGTA 727
Db 1663 CTTCATGAGGAGGAGGAGTGGGATCTACACCAATGTTTAAATATGATCTGATGTA 1722
Qy 728 GAACTGCTAAGAGACAGTGAAGACCTTCTCTCTGCAATTCAGCTGGCTTGCCA 787
Db 1723 GAACTGCTAAGAGACAGTGAAGACCTTCTCTCTGCAATTCAGCTGGCTTGCCA 1782
Qy 788 TGGAATATACAGACAGATTAATTTCCCTCTATTCAAAATTAATTTCCAAATGAAA 844
Db 1783 TGGAATATACAGACAGATTAATTTCCCTCTATTCAAAATTAATTTCCAAATGAAA 1839

RESULT 13
AR240729 AR240729 705 bp DNA 1linear PAT 20-DEC-2002
LOCUS Sequence 3 from patent US 6468776.
DEFINITION AR240729
ACCESSION AR240729
VERSION AR240729.1 GI:27285917
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 705)
AUTHORS Konklin,D.C.
TITLE Human serine protease
JOURNAL Patent: US 6468776-A 3 22-OCT-2002;
FEATURES
source Location/Qualifiers
1..705
/organism="unknown"

BASE COUNT 133 a 68 c 116 g 102 t 286 others
ORIGIN
Query Match 59.0%; Score 511.2; DB 6; Length 705;
Best Local Similarity 59.3%; Pred. No. 5,4e-133;
Matches 418; Conservative 163; Mismatches 124; Indels 0; Gaps 0;

Qy 42 ATGAATATGTCCTTCTATTTGGGTGTCCTGCTGGAGCAATTTTCTTGTGCTGACTATCT 101
Db 1 ATGAATATGTCCTTCTATTTGGGTGTCCTGCTGGAGCAATTTTCTTGTGCTGACTATCT 101
Qy 102 GTTCAGAAAGAACACCTGCTCTCTATTTGGTGTGACTCAAGTCTCACTTCAACCCCTGT 161
Db 61 GTTCAGAAAGAACACCTGCTCTCTATTTGGTGTGACTCAAGTCTCACTTCAACCCCTGT 120
Qy 162 GTGGGGCTCCTCAACCAACAGCTGGGTGCTGGCCCCAGCTCACTGCTATTTACCAAT 221
Db 121 GTGGGGCTCCTCAACCAACAGCTGGGTGCTGGCCCCAGCTCACTGCTATTTACCAAT 180
Qy 222 CTGAAGATGATGCTGGGAAATTTCAAGACAGAGTCAAGCGTACTGAACAGCAAT 281
Db 181 CTGAAGATGATGCTGGGAAATTTCAAGACAGAGTCAAGCGTACTGAACAGCAAT 240
Qy 282 AACCCATTGATGCTGGGAAATTTCAAGACAGAGTCAAGCGTACTGAACAGCAAT 341
Db 241 AACCCATTGATGCTGGGAAATTTCAAGACAGAGTCAAGCGTACTGAACAGCAAT 300
Qy 342 ATGCTCATCAGGCTGCTAAGCTGCTGACATCTCAATCCCAAGTCCAGCCCTTCCCTC 401
Db 401 ATGCTCATCAGGCTGCTAAGCTGCTGACATCTCAATCCCAAGTCCAGCCCTTCCCTC 461

Db 301 ATGTTNATHAARYTNGCNAARCCNGATNTAAYCCNAARGTNCARCCNTNACNTYN 360
Qy 402 GCCACCAACATGTCAGGCGAGGCACTGTCTGATCTACGATTTGGACTGGAGCA 461
Db 361 GCCACCAACATGTCAGGCGAGGCACTGTCTGATCTACGATTTGGACTGGAGCA 420
Qy 462 GAAACAGTGGCGGACACCCCTGACTGCGGAGAACCTGAGAGCCCCGTGATGTCAT 521
Db 421 GAAACAGTGGCGGACACCCCTGACTGCGGAGAACCTGAGAGCCCCGTGATGTCAT 480
Qy 522 CGAATATGCCAAAAACAGACAGGAAAGCCACAGAAATTCCTATGTCGAAATTT 581
Db 481 CGAATATGCCAAAAACAGACAGGAAAGCCACAGAAATTCCTATGTCGAAATTT 540
Qy 582 GTGAAGATTTTCAAGCGGAAATTTTGGGAGGTGGCGCTGCTACCTCATCTGCAAGAC 641
Db 541 GTGAAGATTTTCAAGCGGAAATTTTGGGAGGTGGCGCTGCTACCTCATCTGCAAGAC 600
Qy 642 AAGCTCAGGAAATCAGGTGGGCACTTCATGAGAGGAGCGTCCGATCTACACCAAT 701
Db 601 AAGCTCAGGAAATCAGGTGGGCACTTCATGAGAGGAGCGTCCGATCTACACCAAT 660
Qy 702 GTTACAAATATGATCTCTGATTTGAGAACTGCTAAGACAG 746
Db 661 GTTACAAATATGATCTCTGATTTGAGAACTGCTAAGACAG 705

RESULT 14
AX318281 AX318281 705 bp DNA 1linear PAT 14-DEC-2001
LOCUS Sequence 3 from Patent WO0166771.
DEFINITION AX318281
ACCESSION AX318281
VERSION AX318281.1 GI:1790931
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Konklin,D.C.
TITLE Human serine protease
JOURNAL Patent: WO 0166771-A 3 13-SEP-2001;
FEATURES
source Location/Qualifiers
1..705
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="This degenerate nucleotide sequence encodes the amino acid sequence of SEQ ID NO:2."

BASE COUNT 133 a 68 c 116 g 102 t 286 others
ORIGIN
Query Match 59.0%; Score 511.2; DB 6; Length 705;
Best Local Similarity 59.3%; Pred. No. 5,4e-133;
Matches 418; Conservative 163; Mismatches 124; Indels 0; Gaps 0;

Qy 42 ATGAATATGTCCTTCTATTTGGGTGTCCTGCTGGAGCAATTTTCTTGTGCTGACTATCT 101
Db 1 ATGAATATGTCCTTCTATTTGGGTGTCCTGCTGGAGCAATTTTCTTGTGCTGACTATCT 101
Qy 102 GTTCAGAAAGAACACCTGCTCTCTATTTGGTGTGACTCAAGTCTCACTTCAACCCCTGT 161
Db 61 GTTCAGAAAGAACACCTGCTCTCTATTTGGTGTGACTCAAGTCTCACTTCAACCCCTGT 120
Qy 162 GTGGGGCTCCTCAACCAACAGCTGGGTGCTGGCCCCAGCTCACTGCTATTTACCAAT 221
Db 121 GTGGGGCTCCTCAACCAACAGCTGGGTGCTGGCCCCAGCTCACTGCTATTTACCAAT 180
Qy 222 CTGAAGATGATGCTGGGAAATTTCAAGACAGAGTCAAGCGTACTGAACAGCAAT 281
Db 181 CTGAAGATGATGCTGGGAAATTTCAAGACAGAGTCAAGCGTACTGAACAGCAAT 240
Qy 282 AACCCATTGATGCTGGGAAATTTCAAGACAGAGTCAAGCGTACTGAACAGCAAT 341
Db 341 AACCCATTGATGCTGGGAAATTTCAAGACAGAGTCAAGCGTACTGAACAGCAAT 301

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2003, 22:59:16 ; Search time 326 Seconds
(without alignments)
7170.900 Million cell updates/sec

Title: US-10-037-270-482

Perfect score: 866
Sequence: 1 ggcactactccctgagctca.....tgggaatgacatcactagt 866

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	866	100.0	866	22	AA158594 Human polynucleotide
2	867.8	99.6	1060	24	ABK15700 Human CDNA encodin
3	867.8	99.1	890	24	ABK12989 Human protease PR
4	850.8	98.2	865	22	AA12970 Human Ztryp3 serin
5	850.8	98.2	865	25	ABQ77082 Human serine prote
6	850.2	98.2	867	22	AA160380 Human polynucleoti
7	834	96.3	865	24	ABK30293 Human G-protein-co
8	730	84.3	730	24	ABA04594 MO16a coding seque

9	728.4	84.1	730	24	ABA04595
10	706.4	81.6	708	22	AA544567
11	703.8	81.3	721	24	ABA04647
12	606.4	70.0	798	24	ABK11786
13	511.4	59.1	705	22	AA512971
14	511.2	59.0	705	25	ABQ77083
15	246	28.4	379	21	AA24057
16	154.4	17.8	853	24	AB235087
17	154	17.8	850	24	ABL65438
18	153.6	17.7	2342	25	ACC64648
19	153.2	17.7	802	24	AB235098
20	153.2	17.7	954	20	ABK44208
21	153.2	17.7	802	24	AA277512
22	152.8	17.6	741	9	AA181633
23	152.8	17.6	744	16	AA103999
24	152.8	17.6	744	16	AA104000
25	152.8	17.6	744	16	AA104001
26	152.8	17.6	790	19	AAV24548
27	151.2	17.5	1047	22	AAH57534
28	145	16.7	825	21	AA292585
29	143.6	16.6	1479	22	AAH57533
30	142	16.4	699	22	AA111003
31	142	16.4	701	20	AA23298
32	139.6	16.1	897	18	AA149878
33	132.8	15.3	522	24	ABK44151
34	132.8	15.3	534	24	ABK44157
35	132.8	15.3	566	24	ABK44158
36	132.8	15.3	567	24	ABK44202
37	132.8	15.3	572	24	ABK44081
38	131.8	15.2	592	24	AAV86442
39	131.8	15.2	566	24	ABK44159
40	131	15.1	556	24	ABK44146
41	130.6	15.1	565	24	ABK44152
42	130.4	15.1	566	24	ABK44160
43	130.2	15.0	560	24	ABK44143
44	125.2	14.5	701	15	AA063795
45	123.6	14.3	683	15	AA063794

ALIGNMENTS

RESULT 1
AA158594
ID AA158594 standard; CDNA, 866 BP.
XX
AC AA158594;
XX
DT 22-OCT-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 797.
DE
XX
XX Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; rheumatic;
KW amyotrophic lateral sclerosis; Shy-Dreger Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW Leukemia; ss.
XX
XX Homo sapiens.
OS
XX
XX MO200153312-A1.
PN
XX
XX 26-JUL-2001.
PD
XX
XX 26-DEC-2000; 2000WC-US34263.
PF
XX
XX 21-JAN-2000; 2000US-0488725.
PR
XX 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI: 2001-442253/47.
 DR P-PSDB; AAM39438.
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 XX such as central nervous system injuries -
 XX
 PS Claim 1; SEQ ID NO 797; 10078bp; English.
 CC The invention relates to human nucleic acids (AA15798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemoreactive/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 CC
 SQ Sequence 866 BP; 233 A; 227 C; 197 G; 209 T; 0 other;

Query Match 100.0%; Score 866; DB 22; Length 866;
 Best Local Similarity 100.0%; Pred. No. 7.2e-243;
 Matches 866; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACTTACTTCCTGAGCTAAAGGGGGAAGAGCTGATCCATGAAATATGCTCTTATTT 60
 DB 1 GGCACTTACTTCCTGAGCTAAAGGGGGAAGAGCTGATCCATGAAATATGCTCTTATTT 60
 QY 61 TGGGTGCTCGCTGGGAGATTTTCTTGCTGATCATCTGTTTCAAGAAAGAACCCCTG 120
 DB 61 TGGGTGCTCGCTGGGAGATTTTCTTGCTGATCATCTGTTTCAAGAAAGAACCCCTG 120
 QY 121 CTCCTATTTGTGTATCTCAAGTCTCACTTCAACCCCTGTGTGGGCTCTCATCAAC 180
 DB 121 CTCCTATTTGTGTATCTCAAGTCTCACTTCAACCCCTGTGTGGGCTCTCATCAAC 180
 QY 181 CGAGCTGGGTGCTGGCCCCAGCTCACTGCTATTTTACCAAAATCTGAAGTATGCTGGAA 240
 DB 181 CGAGCTGGGTGCTGGCCCCAGCTCACTGCTATTTTACCAAAATCTGAAGTATGCTGGAA 240
 QY 241 ATTTCAGAGCAGAGTGAAGAGCGTACTGAACAGAAATTAACCCCATCAATGCTCC 300
 DB 241 ATTTCAGAGCAGAGTGAAGAGCGTACTGAACAGAAATTAACCCCATCAATGCTCC 300
 QY 301 GCTACTGGAAGTCAAGTCAATGAGCCCAAGATGAGCTCATGCTCATCAAGTGGCTA 360
 DB 301 GCTACTGGAAGTCAAGTCAATGAGCCCAAGATGAGCTCATGCTCATCAAGTGGCTA 360
 QY 361 AGCCTGCGATGCTCAATCCAAAGTCAGCCCTTCCCTGCGACCAACCAATGTCAGGC 420
 DB 361 AGCCTGCGATGCTCAATCCAAAGTCAGCCCTTCCCTGCGACCAACCAATGTCAGGC 420
 QY 421 CAGGCACTGCTGCTCACTCTCAAGTTTGAAGTGAAGCCAAAGAAAGTGGCGAACC 480
 DB 421 CAGGCACTGCTGCTCACTCTCAAGTTTGAAGTGAAGCCAAAGAAAGTGGCGAACC 480
 QY 481 CTGACTTGCGGAGAACCTGGAAGGCCCGCGTATGTCTGATCGAATGCCAAAAACAG 540

DB 481 CTGACTTGCGGAGAACCTGGAAGGCCCGCGTATGTCTGATCGAATGCCAAAAACAG 540
 QY 541 AACAAAGAAAAGCCACAGAAATTCCTTATGTGTGAATTTGGAAAGTATTCAGCCGA 600
 DB 541 AACAAAGAAAAGCCACAGAAATTCCTTATGTGTGAATTTGGAAAGTATTCAGCCGA 600
 QY 601 TTTTGGGAGGAGTGGCCGTTGCTACTGTCATCTGCAAGACAGCTCCAGGAAATCGAG 660
 DB 601 TTTTGGGAGGAGTGGCCGTTGCTACTGTCATCTGCAAGACAGCTCCAGGAAATCGAG 660
 QY 661 TGGGAGCACTTCATGAGAGGGGAGCTGCGCATCTACACCAATGTTTCAAAATATGATCT 720
 DB 661 TGGGAGCACTTCATGAGAGGGGAGCTGCGCATCTACACCAATGTTTCAAAATATGATCT 720
 QY 721 GATTTGAGAACCTGCTAAGACAAATGAGACCCCTCTCCCTGATTCCTCACTGGC 780
 DB 721 GATTTGAGAACCTGCTAAGACAAATGAGACCCCTCTCCCTGATTCCTCACTGGC 780
 QY 781 TCTGCCATGAGACTATACAGACAGATATTTTCCCTATTCAAAATTAATCTCCAAATG 840
 DB 781 TCTGCCATGAGACTATACAGACAGATATTTTCCCTATTCAAAATTAATCTCCAAATG 840
 QY 841 AAAATTTGGGAATGTAGCATTAAGT 866
 DB 841 AAAATTTGGGAATGTAGCATTAAGT 866

RESULT 2
 ABK15700
 ID ABK15700 standard; cDNA; 1060 BP.
 XX
 AC ABK15700;
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DE Human cDNA encoding novel serine protease 14087.
 XX
 KW Human; ss; gene; serine protease; 14087; cellular proliferative disorder;
 KW cancer; carcinoma; leukaemia; brain disorder; meningitis; AIDS;
 KW acquired immunodeficiency syndrome-associated myopathy; heart disorder;
 KW multiple sclerosis; Alzheimer's disease; Parkinson's disease;
 KW ischemic heart disease; cardiac hypertrophy; myocardial infarction;
 KW atherosclerosis; hypertension; platelet disorder; osteoporosis;
 KW osteomalacia; osteopenia; haematopoietic disorder; diabetes mellitus;
 KW rheumatoid arthritis; liver disorder; viral disease; metabolic disorder;
 KW obesity; anorexia; cachexia.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH CDS 164..871
 FT /*tag= a
 FT /product= "Serine protease 14087"
 FT sig_peptide 164..223
 FT /*tag= b
 FT mat_peptide 224..868
 FT /*tag= c
 XX
 PN MO200206331-A2.
 PD 24-JAN-2002.
 XX
 PR 18-JUL-2001; 2001MO-US23184.
 XX
 PR 18-JUL-2000; 2000US-219022P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Meyers R;
 XX
 DR WPI: 2002-188537/24.
 DR P-PSDB; AAU76316.

XX New serine protease polypeptide for diagnosing and treating cellular
 PT proliferation, brain, heart, bone, liver or metabolic disorders, viral
 PT diseases and for identifying modulators

XX Claim 1; Fig 1; 105pp; English.

XX The invention relates to an isolated serine protease polypeptide, 14087,
 CC or its naturally occurring allelic variant or fragment (and their
 CC encoding polynucleotides). Also included are a host cell containing the
 CC polynucleotide, an anti-14087 antibody and a modulator of the activity
 CC of 14087. 14087 is useful for identifying a compound which binds to or
 CC modulates the activity of 14087. Fragments comprising at least 25
 CC contiguous nucleotides of a nucleic acid encoding 14087 are useful
 CC as hybridization probes or primers for identifying a nucleic acid
 CC molecule associated with a disorder, and a subject having disorder or
 CC at risk for developing a disorder. Assaying the ability of the compound
 CC to modulate 14087 nucleic acid expression or 14087 polypeptide activity
 CC is useful for identifying a compound capable of treating disorder
 CC characterized by aberrant 14087 nucleic acid expression or polypeptide
 CC activity. 14087 proteins are useful for modulating cellular
 CC proliferation, differentiation, tumorigenesis, modulating an immune
 CC response, catalysing muscle-related reactions and modulating
 CC proteolysis of protein substrates. 14087 protein is useful in treating
 CC various disorders including cellular proliferative and/or
 CC differentiative disorders (e.g. cancer, carcinoma, leukemia or
 CC haematopoietic neoplastic disorder), brain (e.g. meningitis, acquired
 CC immunodeficiency syndrome (AIDS)-associated myopathy, multiple
 CC sclerosis, Alzheimer's disease, Parkinson's disease), heart (e.g.
 CC ischaemic heart disease, cardiac hypertrophy, myocardial
 CC infarction), blood vessel (e.g. atherosclerosis, hypertension, Kawasaki
 CC syndrome), and platelet disorders, as well as disorders associated with
 CC bone metabolism (e.g. osteoporosis, osteomalacia, osteopenia, tropical
 CC sprue, rickets), haematopoietic (e.g. diabetes mellitus, rheumatoid
 CC arthritis, autoimmune thyroiditis), liver disorders, viral diseases and
 CC pain or metabolic disorders (e.g. obesity, anorexia, cachexia) and
 CC many other diseases and disorders given in the specification. The
 CC present sequence is the cDNA encoding serine protease 14087.

XX Sequence 1060 BP, 309 A, 259 C, 237 G, 251 T, 4 other;

XX Query Match 99.6%; Score 862.8; DB 24; Length 1060;
 XX Best Local Similarity 99.8%; Pred. No. 6,9e-242;
 XX Matches 864; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGCACCTTACTCCCTGAGCTAAGGGGGAAGAGCTGATCACCATGAATAATGCTTCTATT 60
 DB 123 GACACTTACTCCCTGAGCTAAGGGGGAAGAGCTGATCACCATGAATAATGCTTCTATT 182
 QY 61 TGGGTGCTCCCTGGGAGCAATTTTCTTCTGATCATCTGTTCAAGAAAGAACCTTG 120
 DB 183 TGGGTGCTCCCTGGGAGCAATTTTCTTCTGATCATCTGTTCAAGAAAGAACCTTG 242
 QY 121 CTCCTTATTTGATGATCACTCAAGCTCACTTCAACCCCTGTGGGCGTCTCATCAAC 180
 DB 243 CTCCTTATTTGATGATCACTCAAGCTCACTTCAACCCCTGTGGGCGTCTCATCAAC 302
 QY 181 CCAGCTGGGTGCTGGCCCCAGCTCACTGCTATTATTCGAATCTGAAGATGATCTGGAA 240
 DB 303 CCAGCTGGGTGCTGGCCCCAGCTCACTGCTATTATTCGAATCTGAAGATGATCTGGAA 362
 QY 241 ATTTCAAGAGCAAGCTCAAGAGGCTATCTGAACAGACATTAACCCCATTCAGATGCTC 300
 DB 363 ATTTCAAGAGCAAGCTCAAGAGGCTATCTGAACAGACATTAACCCCATTCAGATGCTC 422
 QY 301 GCTACTGAAGCTACAGTATAGGCCCCACAGAGATGACTCATGCTCATCAAGCTGACTA 360
 DB 423 GCTACTGAAGCTACAGTATAGGCCCCACAGAGATGACTCATGCTCATCAAGCTGACTA 482
 QY 361 AGCTTGCATGCTCAATCCCAAGATCAGGCCCTTCCCTCGGCACACCAATGTGAGG 420
 DB 483 AGCTTGCATGCTCAATCCCAAGATCAGGCCCTTCCCTCGGCACACCAATGTGAGG 542

QY 421 CAGGACCTGTCTGTACTCTCAGGTTTGAGCTGAGCCAGAAAAACAGTGGCCGACACC 480
 DB 543 CAGGACCTGTCTGTACTCTCAGGTTTGAGCTGAGCCAGAAAAACAGTGGCCGACACC 602
 QY 481 CTGACTTGGGCGAAGAACTGGAGGCCCGCTGATGTCTGATGAGAAATGCCAAAAACAG 540
 DB 603 CTGACTTGGGCGAAGAACTGGAGGCCCGCTGATGTCTGATGAGAAATGCCAAAAACAG 662
 QY 541 AACAGAGAAAAAGCCACAGAGAAATTCCTTATGTGTGAAATTTGAAAGATTCAGCCGAA 600
 DB 663 AACAGAGAAAAAGCCACAGAGAAATTCCTTATGTGTGAAATTTGAAAGATTCAGCCGAA 722
 QY 601 TTTTGGGAGGTGGCCGCTTGTCTATCTCATCTGCAAGACAACTCCAGGAAATGAGG 660
 DB 723 TTTTGGGAGGTGGCCGCTTGTCTATCTCATCTGCAAGACAACTCCAGGAAATGAGG 782
 QY 661 TGGGGCATCTCATGGAGGGGACGTGGGATCTACACCAATGTTTCAAAATGTATCTCT 720
 DB 783 TGGGGCATCTCATGGAGGGGACGTGGGATCTACACCAATGTTTCAAAATGTATCTCT 842
 QY 721 GGATTGAGAACCTGCTAAGACAAAGTGAACCTTCTCCCTGCAATTCACACTGGC 780
 DB 843 GGATTGAGAACCTGCTAAGACAAAGTGAACCTTCTCCCTGCAATTCACACTGGC 902
 QY 781 TCTGCATGACTATACAGACAGATTAATTTTCTCTTATTCAAAATPAAATCTCCAAATG 840
 DB 903 TCTGCATGACTATACAGACAGATTAATTTTCTCTTATTCAAAATPAAATCTCCAAATG 962
 QY 841 AAAATTTGGGAATGTAGCATCTACTAGT 866
 DB 963 AAAATTTGGGAATGTAGCATCTACTAGT 988

RESULT 3
 ABK12889
 ID ABK12889 standard; cDNA; 890 BP.
 XX
 AC ABK12889;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Human protease PRS-6 cDNA sequence.
 XX
 XX Human; protease; PRS; gastrointestinal; Crohn's disease; cancer;
 KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
 KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
 KW cell proliferative disorder; developmental disorder; epilepsy;
 KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
 KW reproductive disorder; endometriosis; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 63..770
 FT /*tag= a
 FT /product= "Human protease PRS-6"
 FT
 PN WO200198468-A2.
 PD 27-DEC-2001.
 XX
 PF 13-JUN-2001; 2001WO-US19178.
 XX
 XX 16-JUN-2000; 2000US-212336P.
 PR 22-JUN-2000; 2000US-213955P.
 PR 29-JUN-2000; 2000US-215395P.
 PR 07-JUL-2000; 2000US-216821P.
 PR 14-JUL-2000; 2000US-218946P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 PA Yue H, Elliott VS, Gandhi AR, Lai P, Au-Young J, Tribouley CM;
 PI Deleane AM, Baughn MR, Nguyen DB, Lee EA, Hatfield A, Khan FA;

PI Walia NK, Yao MG, Lu DM, Patterson C, Tang YT, Walsh RT,
PI Aizimel Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L,
PI Kallik DA;
XX
DR WPI: 2002-090437/12.
DR P-PSDB; AAU074746.
XX
PT Twenty one human proteases (referred to as PRS-1 to PRS-21), useful
PT in the diagnosis, treatment and prevention of gastrointestinal (e.g.
PT gastritis), cardiovascular (e.g. atherosclerosis) and cell
PT proliferative (e.g. cancer) disorders -
XX
PS Claim 5; Page 164; 177p; English.
XX
CC The present invention relates to twenty one human proteases,
CC referred to as PRS-1 to PRS-21. The PRS polynucleotides and
CC polypeptides of the invention are useful in the diagnosis, treatment and
CC prevention of gastrointestinal e.g. gastritis, esophageal carcinoma and
CC Crohn's disease, cardiovascular e.g. atherosclerosis, hypertension and
CC myocardial infarction, autoimmune/inflammatory e.g. acquired
CC immunodeficiency syndrome (AIDS), allergies and osteoarthritis, cell
CC proliferative e.g. cancer, developmental e.g. Duchenne and Becker
CC muscular dystrophy, epithelial e.g. dermatitis, neurological e.g.
CC epilepsy and Alzheimer's disease and reproductive e.g. infertility and
CC endometriosis disorders. Numerous other examples of each disorder are
CC given in the specification. The present nucleic acid sequence encodes
CC the human protease PRS-6 protein of the invention.
XX
SQ Sequence 890 BP; 249 A; 230 C; 203 G; 208 T; 0 other:
Query Match 99.1%; Score 857.8; DB 24; Length 890;
Best Local Similarity 99.8%; Pred. No. 1.8e-240;
Matches 859; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGCACTTACTCTCCGAGCTAAGGGGGAAGAGCTGATCACCAGAAATATGCTTCTTATT 60
DB 22 GGCACTTACTCTCCGAGCTAAGGGGGAAGAGCTGATCACCAGAAATATGCTTCTTATT 81
QY 61 TGGGTGCTCGGCTGGGAGCAATTTTCTTGTGCTGATCATCTGTTCAAGAAAGACCTG 120
DB 82 TGGGTGCTCGGCTGGGAGCAATTTTCTTGTGCTGATCATCTGTTCAAGAAAGACCTG 141
QY 121 CTCCTATTTGTGTACTCAAGTCTCACTTCAACCCCTGTGTGGGCTCTCATCAAC 180
DB 142 CTCCTATTTGTGTACTCAAGTCTCACTTCAACCCCTGTGTGGGCTCTCATCAAC 201
QY 181 CGAGCTGGGTGCTGGCCCGAGCTCACTGATTTTCAAAATCGAAAGTATGCTGGAA 240
DB 202 CGAGCTGGGTGCTGGCCCGAGCTCACTGATTTTCAAAATCGAAAGTATGCTGGAA 261
QY 241 ATTTCAAGACAGAGTCAAGAGGCTAAGTCAAGCAATTAACCCGATTCAGATGCTGC 300
DB 262 ATTTCAAGACAGAGTCAAGAGGCTAAGTCAAGCAATTAACCCGATTCAGATGCTGC 321
QY 301 GCTACTGGAATCACTAGTATAGCCGCCACAGATGATCTCATGATGCTGCTA 360
DB 322 GCTACTGGAATCACTAGTATAGCCGCCACAGATGATCTCATGATGCTGCTA 381
QY 361 AGCTGCGATGCTCATATCCAAAGTCAAGCCCTTCCCTGCGACCAACAATGTCAGGC 420
DB 382 AGCTGCGATGCTCATATCCAAAGTCAAGCCCTTCCCTGCGACCAACAATGTCAGGC 441
QY 421 CAGGCACTGTCTGTCTACTCTCAAGGTTTGAAGTGGAGCCAAAGAAAGTGGCCGAC 480
DB 442 CAGGCACTGTCTGTCTACTCTCAAGGTTTGAAGTGGAGCCAAAGAAAGTGGCCGAC 501
QY 481 CTGACTTGGCGAGAACTGAGGCCCCGATGATGCTGATTCGAGATGCAAAAAACAG 540
DB 502 CTGACTTGGCGAGAACTGAGGCCCCGATGATGCTGATTCGAGATGCAAAAAACAG 561
QY 541 AACAGGAAAAAGCCACAGGAATTCCTTATGTGTGAATTTGTGAAGTATTCAGCCGAA 600
DB 562 AACAGGAAAAAGCCACAGGAATTCCTTATGTGTGAATTTGTGAAGTATTCAGCCGAA 621

QY 601 TTTTGGGAGAGTGGCCGTTGCTACTGTCATCTGCAAGACAGCTCCAGGAATCGAG 660
DB 622 TTTTGGGAGAGTGGCCGTTGCTACTGTCATCTGCAAGACAGCTCCAGGAATCGAG 681
QY 661 TGGGGCACTTCATGAGAGGGGAGCTGGGCACTCAACCAATGTTTCAAAATATGATCT 720
DB 682 TGGGGCACTTCATGAGAGGGGAGCTGGGCACTCAACCAATGTTTCAAAATATGATCT 741
QY 721 GATTTAGAACACTGCTAAGAGCAAGTGAAGACCTTCTCCCTGCAATTCACCTGGC 780
DB 742 GATTTAGAACACTGCTAAGAGCAAGTGAAGACCTTCTCCCTGCAATTCACCTGGC 801
QY 781 TCGCCATGAGCTATCAAGCAGATATTTTCCCTATTCAAATTAATCTCAATG 840
DB 802 TCGCCATGAGCTATCAAGCAGATATTTTCCCTATTCAAATTAATCTCAATG 861
QY 841 AAAATTTGGGAATGTACATA 861
DB 862 AAAATTTGGGAATGTACATA 882
RESULT 4
ID AAS12970 standard; DNA; 865 BP.
AC AAS12970;
XX
XX 18-DEC-2001 (first entry)
XX
XX Human Ztryp3 serine protease DNA.
DE
XX Human: Ztryp3; serine protease; asthma; vascular function; inflammation;
KW gene therapy; stroke; testicular function; spermatogenesis; hemostatic;
KW mass spectrometry; circular dichroism; X-ray crystallography; ds;
KW nuclear magnetic resonance spectroscopy; antiasthmatic; antiinflammatory;
KW cerebroprotective.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 42..749
FT /*tag= a
FT /product= "Human Ztryp3"
FT sig_peptide 42..98
FT /*tag= b
FT mat_peptide 99..746
FT /*tag= c
FT /product= "Mature human Ztryp3"
XX
XX W0200166771-A2.
XX
XX 13-SEP-2001.
PD
XX 28-FEB-2001; 2001WO-US06432.
PF
XX 03-MAR-2000; 2000US-0518387.
PR
XX (ZYMO) ZYMOGENETICS INC.
PA
XX Conklin DC;
PI
XX WPI: 2001-589946/66.
DR P-PSDB; AAU07697.
XX
XX Novel Ztryp3 polypeptides and polynucleotides useful in the treatment
PT of asthma, vascular disorders including stroke, inflammation and
PT testicular function -
XX
XX Claim 4; Page 79-81; 82p; English.
PS
XX The invention relates to an isolated human Ztryp3 polypeptide, a member
CC of the serine protease family. Ztryp3 polypeptides and their associated

CC polynucleotides are useful in diagnosis, therapy and industry and are
CC used as targets for identifying modulators, preferably inhibitors of
CC serine protease activity. The sequences are useful in the treatment of
CC asthma, vascular function such as stroke, inflammation and testicular
CC function (by modulating spermatogenesis). Ztryp3 proteins can be used for
CC identifying peptide cleavage sites and for coupling amino and carboxy
CC terminal tags. The polypeptides are also useful to teach analytical
CC skills such as mass spectrometry, circular dichroism, X-ray
CC crystallography and nuclear magnetic resonance spectroscopy. This
CC sequence represents DNA encoding human Ztryp3.

CC
XX
SQ Sequence 865 BP; 235 A; 226 C; 196 G; 208 T; 0 other;

Query Match 98.2%; Score 850.8; DB 22; Length 865;
Best Local Similarity 99.7%; Pred. No. 2e-238;
Matches 865; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 GGCACTTACTCCCTGAGCTAAGGGGGAAGAGCTGATCACCATGAATATGTTCTATT 60
DB 1 GGCACTTACTCCCTGAGCTAAGGGGGAAGAGCTGATCACCATGAATATGTTCTATT 60
QY 61 TGGGTGTCCTCGCTGGGACATTTTCTTCTGCTGATCATCTGTTGAGAAAGAGCCCTG 120
DB 61 TGGGTGTCCTCGCTGGGACATTTTCTTCTGCTGATCATCTGTTGAGAAAGAGCCCTG 120
QY 121 CTCCTATTGTTGTTGCTGAGCTCAAGCTCACTTCAACCCCTGTTGGGGGCTCTCATCAAC 180
DB 121 CTCCTATTGTTGTTGCTGAGCTCAAGCTCACTTCAACCCCTGTTGGGGGCTCTCATCAAC 180
QY 181 CCAGCTGGGTGCTGGGCCCCAGCTCACTGATATTACCAATCTGAAAGTANTGCTGGAA 240
DB 181 CCAGCTGGGTGCTGGGCCCCAGCTCACTGATATTACCAATCTGAAAGTANTGCTGGAA 240
QY 241 ATTTCAGAGCAGATGTCAGAGAGGCTACTGAAACAGACATTAACCCCATTCAGATGCTC 300
DB 241 ATTTCAGAGCAGATGTCAGAGAGGCTACTGAAACAGACATTAACCCCATTCAGATGCTC 300
QY 301 GCTACTGGAACCTACAGTCAATGAGGCCCAAGAGTGAATGCTCATCAAGCTGGCTA 360
DB 301 GCTACTGGAACCTACAGTCAATGAGGCCCAAGAGTGAATGCTCATCAAGCTGGCTA 360
QY 361 AGCCTGCATGCTCAATCCCAAAAGTCCAGCCCTTACCCTGGCACACCAATGTCAGGC 420
DB 361 AGCCTGCATGCTCAATCCCAAAAGTCCAGCCCTTACCCTGGCACACCAATGTCAGGC 420
QY 421 CAGGCACTGCTCTCTACTCTCAAGGTTGAGCTGAGCCAAAGAAACAGTGGCCGACACC 480
DB 421 CAGGCACTGCTCTCTACTCTCAAGGTTGAGCTGAGCCAAAGAAACAGTGGCCGACACC 480
QY 481 CTGACTTGGGGGAGAACTGAGAGGCCCGCTGATGCTGATGCAAGATGCAAAAAACAG 540
DB 481 CTGACTTGGGGGAGAACTGAGAGGCCCGCTGATGCTGATGCAAGATGCAAAAAACAG 540
QY 541 AACAGAGAAAAAGCCACAGAAATTCCTTATGTTGAAATTTGAAAGTATTCAGCCGAA 600
DB 541 AACAGAGAAAAAGCCACAGAAATTCCTTATGTTGAAATTTGAAAGTATTCAGCCGAA 600
QY 601 TTTTGGGGAGGTGGCCGTTGCTACTGTCATCTGCAAAAGCAAGCTCCAGGGAATCGAG 660
DB 601 TTTTGGGGAGGTGGCCGTTGCTACTGTCATCTGCAAAAGCAAGCTCCAGGGAATCGAG 660
QY 661 TGGGGCACTTCATGAGAGGAGAGAGTGGGCTCTACACCATGTTTACAAATATGATCTCT 720
DB 661 TGGGGCACTTCATGAGAGGAGAGAGTGGGCTCTACACCATGTTTACAAATATGATCTCT 720
QY 721 GGATTGAGAACACTGCTAAGAGACAGACCTTCTCCCTCGATTCGATTCGATGCG 780
DB 721 GGATTGAGAACACTGCTAAGAGACAGACCTTCTCCCTCGATTCGATTCGATTCGATGCG 780
QY 781 TCTGCCATGAGCTATACAGCAGATTAATTTCCCTCTATTCAAATTAATTCGCAATG 840
DB 781 TCTGCCATGAGCTATACAGCAGATTAATTTCCCTCTATTCAAATTAATTCGCAATG 840
QY 840 TCTGCCATGAGCTATACAGCAGATTAATTTCCCTCTATTCAAATTAATTCGCAATG 899
DB 840 TCTGCCATGAGCTATACAGCAGATTAATTTCCCTCTATTCAAATTAATTCGCAATG 899

QY 841 AAAATTGGGAATGTAGCATACTAGT 866
DB 840 AAAATTGGGAATGTAGCAACTAGT 865

RESULT 5
ABQ77082
ID ABQ77082 standard; DNA; 865 BP.
XX
XX ABQ77082;
AC
XX
XX 01-APR-2003 (first entry)
DT
XX
XX Human serine protease Ztryp3 DNA.

DE Human serine protease; Ztryp3; gene; blood coagulation; fibrinolysis;
KW Human; serine protease; Ztryp3; gene; blood coagulation; fibrinolysis;
KW complement activation; fertilization; hormone production; gene therapy;
KW somatic cell; ds.
XX
XX Homo sapiens.
FH
FH Key Location/Qualifiers
FT CDS 42..749
FT /tag= a
FT /product= "Ztryp3"
XX
XX US6468776-B1.
XX
XX 22-OCT-2002.
XX
XX 28-FEB-2001; 2001US-0796110.
XX
XX 03-MAR-2000; 2000US-186623P.
XX
XX (ZYMO) ZYMOGENETICS INC.
XX
XX Conklin DC;
XX
XX WPI; 2003-182091/18.
XX
XX P-PSDB; ABG73741.

PT Novel serine protease nucleic acid molecule useful for producing serine
PT protease polypeptide, and as educational tool in laboratory practicum
PT kits for courses related to genetics and molecular biology
XX
XX
XX Claim 2; Column 55-58; 32pp; English.

XX
XX This invention describes a novel human serine protease, Ztryp3. The
CC serine protease family play a role in carefully controlled processes such
CC as blood coagulation, fibrinolysis, complement activation, fertilization
CC and hormone production. The products of the invention are useful for
CC producing serine protease polypeptides; as probes or primers to clone 5'
CC non-coding regions of Ztryp3 gene; as polymerase chain reaction primers
CC and probes; as educational tool in laboratory practicum kits for courses
CC related to genetics and molecular biology, protein chemistry, and
CC antibody production and analysis; as standards or as unknowns for testing
CC purposes; as an aid to teach a student how to prepare expression
CC constructs for bacterial, viral, or mammalian expression, including
CC fusion constructs, where Ztryp3 is the gene to be expressed; for
CC determining the restriction endonuclease cleavage sites of the
CC polynucleotides; for determining mRNA and DNA localization of Ztryp3
CC polynucleotides in tissues (i.e., by northern and southern blotting as
CC well as polymerase chain reaction); and for identifying related
CC polynucleotides by nucleic acid hybridization. The products of the
CC invention can also be used for determining mutations in the Ztryp3 gene,
CC for determining Ztryp3 gene expression, for examining Ztryp3 gene
CC structure, for detecting and localizing expression of Ztryp3 gene in a
CC biological sample, in vivo diagnosis, in gene therapy, particularly
CC in somatic cell gene therapy, and for producing transgenic mice. This
CC sequence encodes the human serine protease Ztryp3 described in the method
CC of the invention.
XX
XX Sequence 865 BP; 235 A; 226 C; 196 G; 208 T; 0 other;

Query Match 98.2%; Score 850.8; DB 25; Length 865;
Best Local Similarity 99.7%; Pred. No. 2e-238;
Matches 863; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```

QY 1 GGCACCTACTCCCGAGCTAAGGGGGAAGAGCTGATCAACCAATATATCTTCTTATT 60
DB 1 GGCACCTACTCCCGAGCTAAGGGGGAAGAGCTGATCAACCAATATATCTTCTTATT 60
QY 61 TGGGTGCTCTGCTGGGACATTTTCTTCTTCTGATCACTGTTGAGAAAGAACCTTG 120
DB 61 TGGGTGCTCTGCTGGGACATTTTCTTCTTCTGATCACTGTTGAGAAAGAACCTTG 120
QY 121 CTCCCTATTGGTGTACTCAAGTCTCAATCCCTGATGAGGCGTCTCATCAAC 180
DB 121 CTCCCTATTGGTGTACTCAAGTCTCAATCCCTGATGAGGCGTCTCATCAAC 180
QY 181 CCAGCTGGGTGCTGGGCTGCTGATCTGATTTTACCAATCTGAAAGTATGCTGGGAA 240
DB 181 CCAGCTGGGTGCTGGGCTGCTGATCTGATTTTACCAATCTGAAAGTATGCTGGGAA 240
QY 241 ATTTCAAGAGCAGATCAGAGACGGTACTGAAAGACAAATTAACCCATTCAGATGCTC 300
DB 241 ATTTCAAGAGCAGATCAGAGACGGTACTGAAAGACAAATTAACCCATTCAGATGCTC 300
QY 301 GCTACTGGAACCTACAGTCTATAGCGCCCAAGATGATCTGATCAAGCTGCTA 360
DB 301 GCTACTGGAACCTACAGTCTATAGCGCCCAAGATGATCTGATCAAGCTGCTA 360
QY 361 AGCTGCTGATCTGATCAATCCCAAGTCCAGCCCTTCCCTGCGACCAACCAATGTCAGC 420
DB 361 AGCTGCTGATCTGATCAATCCCAAGTCCAGCCCTTCCCTGCGACCAACCAATGTCAGC 420
QY 421 CAGGCACTGCTGCTGATCTGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 480
DB 421 CAGGCACTGCTGCTGATCTGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 480
QY 481 CTGACTTGGCGCAGAACCTGAGGCGCCCGTGAATGTCGATGAGTGAAGTGAAGTGAAGT 540
DB 481 CTGACTTGGCGCAGAACCTGAGGCGCCCGTGAATGTCGATGAGTGAAGTGAAGTGAAGT 540
QY 541 AACAGAGAAAAAGCAGACAGAAATCTTATGATGAAATTTGAAAGTATTCAGCCGAA 600
DB 541 AACAGAGAAAAAGCAGACAGAAATCTTATGATGAAATTTGAAAGTATTCAGCCGAA 600
QY 601 TTTTGGGAGGATGAGGCGCTGCTGATGATGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 660
DB 601 TTTTGGGAGGATGAGGCGCTGCTGATGATGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 660
QY 661 TGGGGCACTTCATGAGGAGGAGCGTCCGCACTTACACCAATGTTTACCAATATGATCTCT 720
DB 661 TGGGGCACTTCATGAGGAGGAGCGTCCGCACTTACACCAATGTTTACCAATATGATCTCT 720
QY 721 GGAATGGAACCTGCTAAGGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 780
DB 721 GGAATGGAACCTGCTAAGGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 780
QY 781 TGTGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
DB 781 TGTGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
QY 841 AAAATTTGGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 865
DB 841 AAAATTTGGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 865

```

RESULT 6
AA160380/c
ID AA160380 standard; cDNA; 867 BP.

XX AC AA160380;
XX DT 22-OCT-2001 (first entry)

```

XX DE Human polynucleotide SEQ ID NO 4369.
XX KW Human, neurotrophic; immunosuppressant; cytostatic; gene therapy; cancer;
XX KW peripheral nervous system; neuropathy; central nervous system; CNS;
XX KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX KW leukaemia; ss.
XX OS Homo sapiens.
XX PN M020015312-A1.
XX PD 26-JUL-2001.
XX PE 26-DEC-2000; 2000MO-US34263.
XX PR 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-0552317.
XX PR 09-JUL-2000; 2000US-0598042.
XX PR 19-JUL-2000; 2000US-0620312.
XX PR 03-AUG-2000; 2000US-0653450.
XX PR 14-SEP-2000; 2000US-0662191.
XX PR 19-OCT-2000; 2000US-0693036.
XX PR 29-NOV-2000; 2000US-0727344.
XX PA (HYSE-) HYSEQ INC.
XX PT Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX PT Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX PT Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX DR WPI; 2001-442253/47.
XX DR P-PSDB; AAM41224.
XX PT Novel nucleic acids and polypeptides, useful for treating disorders
XX PT such as central nervous system injuries -
XX PS Claim 1; SEQ ID NO 4369; 10078pp; English.
XX CC The invention relates to human nucleic acids (AA157798-AA161369) and
XX CC the encoded polypeptides (AAM38642-AAM42213) with neurotrophic,
XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful
XX CC in gene therapy. A composition containing a polypeptide or polynucleotide
XX CC of the invention may be used to treat diseases of the peripheral nervous
XX CC system, such as peripheral nervous injuries, peripheral neuropathy and
XX CC localised neuropathies and central nervous system diseases, such as
XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX CC utilisation of the activities such as: immune system suppression,
XX CC Actin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX CC assays for receptor activity, arthritis and inflammation, leukaemia and
XX CC C.N.S disorders.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification.
XX SQ Sequence 867 BP; 209 A; 197 C; 225 G; 236 T; 0 other;

```

Query Match 98.2%; Score 850.2; DB 22; Length 867;
Best Local Similarity 99.5%; Pred. No. 3e-238;
Matches 863; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```

QY 1 GGCACCTACTCCCGAGCTAAGGGGGAAGAGCTGATCAACCAATATATCTTCTTATT 60
DB 867 GGCACCTACTCCCGAGCTAAGGGGGAAGAGCTGATCAACCAATATATCTTCTTATT 808
QY 61 TGGGTGCTCTGCTGGGACATTTTCTTCTGATCACTGTTGAGAAAGAACCTTG 120
DB 807 TGGGTGCTCTGCTGGGACATTTTCTTCTGATCACTGTTGAGAAAGAACCTTG 748
QY 121 CTCCCTATTGGTGTACTCAAGTCTCAATCCCTGATGAGGCGTCTCATCAAC 180

```

```

Db      747 CTCCTATTGGTGTACTCAAGTCTCACTTCAACCCCTGTGGGCTCTCATCAAC
Qy      181 CGAGCTGGGTGTGGCCCGACCTCACTGCTATTTCACCAATCTGAAGTAGTCTGGAA
Db      687 CGAGCTGGGTGTGGCCCGACCTCACTGCTATTTCACCAATCTGAAGTAGTCTGGAA
Qy      241 ATTTCAAGACAGTCAAGACGCTACTGACAGCAATTAACCCATTCAATGCTCC
Db      627 ATTTCAAGACAGTCAAGACGCTACTGACAGCAATTAACCCATTCAATGCTCC
Qy      301 GCTACTGGAATCTACAGTCAATAGGCCCCACAGATGACTCTGCTCATCACTGGCTA
Db      567 GCTACTGGAATCTACAGTCAATAGGCCCCACAGATGACTCTGCTCATCACTGGCTA
Qy      361 AGCCTGCAATGCTCAATCCCAAGTCC-AGCCCTTCCCTCCGCCACCAATGTCAGG
Db      507 AGCCTGCAATGCTCAATCCCAAGTCCCAAGTCCCAAGTCCCAAGTTCAGG
Qy      420 CCAGGACTGTCTGTCTACTCTCAGGTTTGGACTGGACCAAGAAAAGTGGCCGAC
Db      447 CCAGGACTGTCTGTCTACTCTCAGGTTTGGACTGGACCAAGAAAAGTGGCCGAC
Qy      480 CTTGACTTGGCGGAGAACTTGAGGCCGCCGTGATGTTGATGAGAAATGCCAAAAACA
Db      387 CTTGACTTGGCGGAGAACTTGAGGCCGCCGTGATGTTGATGAGAAATGCCAAAAACA
Qy      540 GAACAAGGAAAAAGCCACAGGAATTCCTTATGTGTGAAATTTGTGAAGTATTCAGCCA
Db      327 GAACAAGGAAAAAGCCACAGGAATTCCTTATGTGTGAAATTTGTGAAGTATTCAGCCA
Qy      600 ATTTTGGGAGGTGGCCGCTTGTCTACTGTCATCTGCAAAAGCAAGCTCCAGGAATCGAG
Db      267 ATTTTGGGAGGTGGCCGCTTGTCTACTGTCATCTGCAAAAGCAAGCTCCAGGAATCGAG
Qy      660 GTGGGCACTTCAATGAGGAGGAGCGTGGCATCTACCAATGTTTAAATATGATCC
Db      207 GTGGGCACTTCAATGAGGAGGAGCGTGGCATCTACCAATGTTTAAATATGATCC
Qy      720 TGGATTGAGAACCTGCTTAAGGACCAAGTGAACCTTCTCCCTGCTGATTCAGTGG
Db      147 TGGATTGAGAACCTGCTTAAGGACCAAGTGAACCTTCTCCCTGCTGATTCAGTGG
Qy      780 CTTGCGCATGAGTATACAGAGATATTTCCCTCTATTCAAAATTAATCTCAAAAT
Db      87 CTTGCGCATGAGTATACAGAGATATTTCCCTCTATTCAAAATTAATCTCAAAAT
Qy      840 GAAAATTTGGGAATGATAGTACTAGT 866
Db      27 GAAAATTTGGGAATGATAGTACTAGT 1

```

RESULT 7
ABK30293/c
ID ABK30293 standard; cDNA; 865 BP.

```

XX      ABK30293;
AC      ABK30293;
DT      23-APR-2002 (first entry)
XX      Human G-protein-coupled protease #63.
DE      Human; aa; gene; G-protein-coupled protease; gene therapy;
XX      transgenic; protease mediated disorder; proliferative disorder;
KM      differential disorder; developmental disorder;
KW      haematopoietic disorder;
XX      Homo sapiens.
XX      US6331427-B1.
PN      18-DEC-2001.
PD
XX

```

```

PF      26-MAR-1999; 99US-0280116.
PR      26-MAR-1999; 99US-0280116.
XX      (MILL-) MILLENNIUM PHARM INC.
XX      Robison KE;
XX      WPI; 2002-129545/17.
DR      New polynucleotides encoding protease homologs of the G-protein-coupled
PT      protease family, useful in identifying agonists and antagonists for
PT      diagnosis and treatment of protease mediated disorders -
PS      Disclosure; Column 109-112; 246pp; English.
XX      The invention relates to an isolated human protease nucleic acid molecule
CC      comprising a nucleotide sequence of 546 base pairs, one of 268 fully
CC      defined in the specification. Also disclosed are production of an
CC      isolated polypeptide encoded by the nucleic acid, comprising introducing
CC      the nucleic acid into a host cell and culturing under conditions to
CC      express the protein from the nucleic acid, use of an antibody to
CC      detect the encoded protein in a sample and to modulate its in vivo
CC      activity, identifying agents that bind to the protein and identification
CC      of a polynucleotide agent that modulates the expression of the nucleic
CC      acid or its complement (i.e. gene therapy). The nucleic acid can be used
CC      to identify an agent that modulates the expression or activity of the
CC      nucleic acid, and can be used to isolate the protein. The nucleic acid
CC      can be used in diagnostic assays for determining nucleic acid expression
CC      as well as activity in the context of a biological sample (e.g., blood,
CC      serum, cells, tissue) to determine whether an individual has a disease or
CC      disorder, or is at risk of developing a disease or disorder, associated
CC      with aberrant expression or activity of the nucleic acid. The nucleic
CC      acid can be used to detect mutations in protease genes and gene
CC      expression products such as mRNA. The nucleic acid can be used as
CC      hybridization probes to detect naturally-occurring genetic mutations in
CC      a protease gene. The nucleic acid can be used in drug screening methods
CC      to identify agonists and antagonists that can be used to diagnose and
CC      treat such protease mediated disorders e.g., proliferative,
CC      differential, developmental or haematopoietic disorders. The nucleic
CC      acid can be used as probes, primers, in biological assays, to determine
CC      patterns of gene expression, to design ribozymes and to construct
CC      transgenic animals. The present sequence represents one of the 268
CC      disclosed human G-protein-coupled protease cDNA sequences.
XX      Sequence 865 BP; 207 A; 197 C; 225 G; 236 T; 0 other;
SQ

```

Query Match 96.3%; Score 834; DB 243; Length 865;
Best Local Similarity 98.7%; Pred. No. 1.7e-233;
Matches 851; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

```

Qy      1 GGCACCTACTCCCTGAGCTAAGGGGGAAGAGCTGATCCATGAAATATGTTCTTATT 60
Db      865 GGCACCTACTCCCTGAGCTAAGGGGGAAGAGCTGATCCATGAAATATGTTCTTATT 806
Qy      61 TGGGTGTCTGCTGGGAGCAATTTTCTTGTGCTGATCATCTGTTTCAAGAAAGACCTTG 120
Db      805 TGGGTGTCTGCTGGGAGCAATTTTCTTGTGCTGATCATCTGTTTCAAGAAAGACCTTG 746
Qy      121 CTCCTAATTTGTTGATCACTCAAGTCTCACTTCAACCCCTGTGGGCGTCTCATCAAC 180
Db      745 CTCCTAATTTGTTGATCACTCAAGTCTCACTTCAACCCCTGTGGGCGTCTCATCAAC 686
Qy      181 CCAGCTGGGTGTGGCCCGACCTCACTGCTATTTCACCAATCTGAAGTAGTCTGGAA 240
Db      685 CCAGCTGGGTGTGGCCCGACCTCACTGCTATTTCACCAATCTGAAGTAGTCTGGAA 626
Qy      241 ATTTCAAGACAGTCAAGACGCTACTGACAGCAATTAACCCATTCAATGCTCC 300
Db      625 ATTTCAAGACAGTCAAGACGCTACTGACAGCAATTAACCCATTCAATGCTCC 566
Qy      301 GCTACTGGAATCTACAGTCAATAGGCCCCACAGATGACTCTGCTCATCACTGGCTA 360

```

```

Db      565  GCTACTGGAAGTACATGATACGCCCCCAAGATGACCTCATCATCAAGCTGCTA 506
Qy      361  AGCTGCGCATGCTCATATCCCAAGTCCAGCCCCCTT-CCCTGCGCACCAACCATGTCAG 419
Db      505  AGCTGCGCATGCTCATATCCCAAGTCCAGCCCCCTTACCCTTCCGACCAACCAATGTCAAG 446
Qy      420  CCAGGCACTGCTGCTCTACTCTCAAGTTTGAGCTGAGCCCAAGAAAACAGTGGCCGACAC 479
Db      445  CCGGGCACTGCTGCTCTACTCTCAAGTTTGAGCTGAGCCCAAGAAAACAGTGGCCGACAC 386
Qy      480  CCTGACTTGGGGGAGAACCTTGAAGGCCCGGTGATGCTGATGAGAAATGCAAAAAACA 539
Db      385  CTTACTTGGGGGAGAACCTTGAAGGCCCGGTGATGCTGATGAGAAATGCAAAAAACA 326
Qy      540  GAACAAGAAAAAAGCCACAGAAATTCCTTATGTGTAATTTGAAAAGTATTCAGCCGA 599
Db      325  GAACAAGAAAAAAGCCACAGAAATTCCTTATGTGTAATTTGAAAAGTATTCAGCCGA 266
Qy      600  ATTTTGGGGAGTGGCCGTTGCTACTGTCATCTGCAAAAGCAAGCTCCAGGGAATCGAG 659
Db      265  ATTTTGGGGAGTGGCCGTTGCTACTGTCATCTGCAAAAGCAAGCTCCAGGGAATCGAG 206
Qy      660  GTGGGGCACTTCATGAGAGGGGAGCTGGGCACTTCAACCAATGTTTCAAAATATGATCC 719
Db      205  GTGGGGCACTTCATGAGAGGGGAGCTGGGCACTTCAACCAATGTTTCAAAATATGATCC 146
Qy      720  TGAATTGAGAACACTGCTTAAGAGCAAGAGACCTTCTCCCTCTGATTCACACTGG 779
Db      145  TGAATTGAGAACACTGCTTAAGAGCAAGAGACCTTCTCCCTCTGATTCACACTGG 86
Qy      780  CTCTGCGATGAGCTATACAGACAGATATTTCCCTCTATTCAAATTAATTCCTCAAT 839
Db      85  CTCTGCGATGAGCTATACAGACAGATATTTCCCTCTATTCAAATTAATTCCTCAAT 26
Qy      840  GAAATTTGGGAATGTGCATTA 861
Db      25  GAAATTTGGGAATGTGCATTA 4

RESULT 8
ABAO4594
ID      ABAO4594 standard; cDNA; 730 BP.
XX
AC      ABAO4594;
XX
DT      21-FEB-2002 (first entry)
XX
DE      MOL6a coding sequence.
XX
KW      MOL; G-coupled protein-receptor; cardiomyopathy; atherosclerosis;
KW      cell signal processing; metabolic disorder; diabetes; cancer;
KW      neurodegenerative disorder; immune disorder; cardiac disorder;
KW      lung disease; autoimmune disease; developmental disorder; anti-diabetic;
KW      Cytoskeletal; Neuroprotective; Antiatherosclerotic; Immunosuppressive;
KW      Gene therapy; Vaccine; antiinflammatory; MOL6a; trypsin-like; ss.
XX
OS      Unidentified.
XX
FH      Key      Location/Qualifiers
FT      CDS      8..715
FT      /*tag= a
FT      /product= "MOL6a"
XX
PN      MO200181578-A2.
XX
PD      01-NOV-2001.
XX
PF      26-APR-2001; 2001WO-US33578.
XX
PR      26-APR-2000; 2000US-200158P.
PR      28-APR-2000; 2000US-200613P.
PR      28-APR-2000; 2000US-200780P.
PR      01-MAY-2000; 2000US-201006P.

```

```

PR      01-MAY-2000; 2000US-201007P.
PR      01-MAY-2000; 2000US-201236P.
PR      01-MAY-2000; 2000US-201238P.
PR      02-MAY-2000; 2000US-201166P.
PR      03-MAY-2000; 2000US-201474P.
PR      03-MAY-2000; 2000US-201508P.
PR      25-JUL-2000; 2000US-220591P.
PR      15-SEP-2000; 2000US-232678P.
PR      22-JAN-2001; 2001US-263217P.
PR      30-JAN-2001; 2001US-265160P.
XX
PA      (CURA-) CURAGEN CORP.
XX
PI      Vernet CAM, Fernandes ER, Gerlach V, Shimkets RA, Mal'yan'kar UM;
PI      Boldog FL, Zernusen BD, Spytek KA, Majumder K, Tchernev VT;
PI      Padiguru M, Patturajan M, Burgess CE, Gangolli EA, Smltson G;
PI      Rastelli L, Macdougall JR, Taupier RJ, Grose WM, Szekeres ES;
PI      Alsobrook JP;
XX
DR      WPI; 2002-049278/06.
DR      P-PSDB; AAM47664.
XX
PT      Novel G-protein coupled receptor-related polypeptides and
PT      polynucleotides for diagnosing, preventing and treating cardiomyopathy,
PT      atherosclerosis, disorders related to cell signal processing and for
PT      identifying modulators -
XX
PS      Claim 8; Page 43-44; 227pp; English.
XX
CC      The present invention relates to novel G-coupled protein-receptor related
CC      proteins and coding sequences (MOLX, where X is a number from 1 to 10,
CC      ABAO4589-ABA04603 and AAM47659-AAM47673). MOLX proteins and coding
CC      sequences are useful for treating or preventing a MOLX-associated
CC      disorder, such as cardiomyopathy, atherosclerosis, disorders related to
CC      cell signal processing and metabolic pathway modulation, diabetes and
CC      cancer. Additionally, MOLX proteins and coding sequences are useful for
CC      preventing and treating a variety of disorders including metabolic
CC      disorders, nutritional oedema, chronic and hereditary pancreatitis,
CC      obesity, infectious disease, anorexia, neurodegenerative disorders,
CC      Alzheimer's disease, Parkinson's disease, stroke, immune disorders,
CC      hematopoietic disorders and various dyslipidaemias, metabolic syndrome X
CC      and wasting disorders associated with chronic diseases and cancers,
CC      cardiac disorders, hypertension, hypercalcaemia, cirrhosis, angiodysplasia
CC      and wound healing, trauma, glomerulonephritis, hyper and hypothyroidism,
CC      multiple sclerosis, lung diseases including asthma, Crohn's disease,
CC      scleroderma, autoimmune diseases, developmental disorders and neural tube
CC      defects. The present sequence is the coding sequence for MOL6a.
CC      MOL6a is a trypsin-like protein.
XX
SQ      Sequence 730 BP; 191 A; 197 C; 170 G; 172 T; 0 other;
XX
Query Match      84.3%; Score 730; DB 24; Length 730;
Best Local Similarity 100.0%; Pred. No. 4,1e-203;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      35  GATCACCATGAATATGCTCTTCTATTGGGTGCTCGTGGAGCATTTTCTTCTGTA 94
Db      1  GATCACCATGAATATGCTCTTCTATTGGGTGCTCGTGGAGCATTTTCTTCTGTA 60
Qy      95  CTCATCTGTTGAGAAAGAACCCCTGCTCCCTATTTGGTGTACTCAAGTTCATTCAA 154
Db      61  CTCATCTGTTGAGAAAGAACCCCTGCTCCCTATTTGGTGTACTCAAGTTCATTCAA 120
Qy      155  CCCCTGTGGGGGCTCTCATCAAAACCAAGTGGGTGCTGGCCCGACCTACGCTATTT 214
Db      121  CCCCTGTGGGGGCTCTCATCAAAACCAAGTGGGTGCTGGCCCGACCTACGCTATTT 180
Qy      215  ACCAAATCTGAAGTATGCTGGGAATTTCAAGACAGAGTCAGACGGTACTGAACA 274
Db      181  ACCAAATCTGAAGTATGCTGGGAATTTCAAGACAGAGTCAGACGGTACTGAACA 240
Qy      275  GACAATTAACCCCATTCAGATCGTCGCTAAGTGAACCTACAGTCATAGCGCCACAGGA 334

```

```

DB      241 GACAAATTAACCCCATTCAGATTCCTCCGCTACTGGAACACTACATCATATAGCGCCACAGAGA 300
QY      335 TGACCTCATGCTCATAGCTGGCTGAAGCTGTCATGCTCAATCCAAATGCCAGCCCT 394
DB      301 TGACCTCATGCTCATAGCTGGCTGAAGCTGTCATGCTCAATCCAAATGCCAGCCCT 360
QY      395 TCCCTCGCCACCAACCAATGTCAGGCGAGGCACTGTCGTCTACTCTCAGGTTGGACTG 454
DB      361 TCCCTCGCCACCAACCAATGTCAGGCGAGGCACTGTCGTCTACTCTCAGGTTGGACTG 420
QY      455 GAGCCAAAGAAACAGTGGCCGACACCTCTGCTGGCGAGAACCTGAGGCCCGCTGANT 514
DB      421 GAGCCAAAGAAACAGTGGCCGACACCTCTGCTGGCGAGAACCTGAGGCCCGCTGANT 480
QY      515 GTCGTGATCGAAGATGCGCAAAAACAGAACAGAAAAAGCACAAGAAATTCCTTAATGTGT 574
DB      481 GTCGTGATCGAAGATGCGCAAAAACAGAACAGAAAAAGCACAAGAAATTCCTTAATGTGT 540
QY      575 GAAATTTGTAAGATGATTCAGCCGAATTTTGGGAGAGTGGCCGTTGCTACTGTCAATCTG 634
DB      541 GAAATTTGTAAGATGATTCAGCCGAATTTTGGGAGAGTGGCCGTTGCTACTGTCAATCTG 600
QY      635 CAAAGCAAGCTCCAGGAAATCGAGGTGGGCACTTCATGAGAGGAGCGCTCGGCACTTA 694
DB      601 CAAAGCAAGCTCCAGGAAATCGAGGTGGGCACTTCATGAGAGGAGCGCTCGGCACTTA 660
QY      695 CACCAATGTTTAAATATGTAATCTGATGAGTGAACACCTGCTAAGGACAAAGTGAACCC 754
DB      661 CACCAATGTTTAAATATGTAATCTGATGAGTGAACACCTGCTAAGGACAAAGTGAACCC 720
QY      755 TACTTCTCCC 764
DB      721 TACTTCTCCC 730

RESULT 9
ABA04595
ID      ABA04595 standard; cDNA; 730 BP.
XX      ABA04595;
AC      21-FEB-2002 (first entry)
XX      MOLdb coding sequence.
DE      MOLdb coding sequence.
XX      MOL; G-coupled protein-receptor; cardiomyopathy; atherosclerosis;
KW      cell signal processing; metabolic disorder; diabetes; cancer;
KW      neurodegenerative disorder; immune disorder; cardiac disorder;
KW      lung disease; autoimmune disease; developmental disorder; anti-diabetic;
KW      cytotoxic; Neuroprotective; Antiatherosclerotic; Immunosuppressive;
KW      Gene therapy; Vaccine; antiinflammatory; MOLdb; trypsin-like; ss.
XX      Unidentified.
OS      Unidentified.
XX      Key CDS      Location/Qualifiers
FH      CDS      8..715
FT      CDS      /tag= a
FT      CDS      /product= "MOLdb"
XX      MO200181578-A2.
XX      01-NOV-2001.
XX      26-APR-2001; 2001MO-US13578.
XX      26-APR-2000; 2000US-200158P.
XX      28-APR-2000; 2000US-200613P.
XX      28-APR-2000; 2000US-200780P.
XX      01-MAY-2000; 2000US-201006P.
XX      01-MAY-2000; 2000US-201007P.
XX      01-MAY-2000; 2000US-201236P.
XX      01-MAY-2000; 2000US-201238P.
XX      02-MAY-2000; 2000US-201186P.

```

```

PR      03-MAY-2000; 2000US-201474P.
PR      03-MAY-2000; 2000US-201508P.
PR      25-JUL-2000; 2000US-220591P.
PR      15-SEP-2000; 2000US-232678P.
PR      22-JAN-2001; 2001US-263217P.
PR      30-JAN-2001; 2001US-265160P.
XX      (CURA-) CURAGEN CORP.
XX      Vernet CM, Fernandes ER, Gerlach V, Shimkete RA, Malpankar UM;
XX      Boldog FL, Zernhsen BD, Spytek KA, Majumder K, Tchernev VT;
XX      Padigar M, Patursajan M, Burgess CE, Gangoli BA, Smithson G;
XX      Rastelli L, Macdougall JR, Taupier RJ, Grose WM, Szekeres ES;
XX      Alsobrook JP;
XX      WPI, 2002-049278/06.
XX      P-PSDB; AAM47665.
XX      Novel G-protein coupled receptor-related polypeptides and
XX      polynucleotides for diagnosing, preventing and treating cardiomyopathy,
XX      atherosclerosis, disorders related to cell signal processing and for
XX      identifying modulators
XX      Claim 8; Page 45; 227p; English.
XX      The present invention relates to novel G-coupled protein-receptor related
XX      proteins and coding sequences (MOLX, where X is a number from 1 to 10,
XX      ABA04589-ABA04603 and AAM47659-AAM47673). MOLX proteins and coding
XX      sequences are useful for treating or preventing a MOLX-associated
XX      disorder, such as cardiomyopathy, atherosclerosis, disorders related to
XX      cell signal processing and metabolic pathway modulation, diabetes and
XX      cancer. Additionally, MOLX proteins and coding sequences are useful for
XX      preventing and treating a variety of disorders including metabolic
XX      disorders, nutritional oedema, chronic and hereditary pancreatitis,
XX      obesity, infectious disease, anorexia, neurodegenerative disorders,
XX      CC Alzheimer's disease, Parkinson's disease, stroke, immune disorders,
XX      CC haematopoietic disorders and various dyslipidaemias, metabolic syndrome X
XX      and wasting disorders associated with chronic diseases and cancers,
XX      CC cardiac disorders, hypertension, hypercalcaemia, cirrhosis, angioneitis
XX      CC and wound healing, trauma, glomerulonephritis, hyper and hypothyroidism,
XX      CC multiple sclerosis, lung diseases including asthma, Crohn's disease,
XX      CC scleroderma, autoimmune diseases, developmental disorders and neural tube
XX      CC defects. The present sequence is the coding sequence for MOLdb.
XX      MOLdb is a trypsin-like protein.
XX      Sequence 730 BP; 191 A; 198 C; 170 G; 171 T; 0 other;
XX      Query Match      84.1%; Score 728.4; DB 24; Length 730;
XX      Best Local Similarity 99.9%; Pred. No. 1.2e-202;
XX      Matches 729; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      35 GATCACCATGAATAATGCTTCTATTGGGTGCTCCTCGTGGACATTTTCTTTGCTGA 94
DB      1 GATCACCATGAATAATGCTTCTATTGGGTGCTCCTCGTGGACATTTTCTTTGCTGA 60
QY      95 CTCATCTGTTCAAGAAAGAACCCCTGCTCTATTGGTGTACTTCAATCTCACTTCAA 154
DB      61 CTCATCTGTTCAAGAAAGAACCCCTGCTCTATTGGTGTACTTCAATCTCACTTCAA 120
QY      155 CCCCTGTGTGGGCGTCTCATCAAAACCCAGCTGGGTGCTGGCCCAAGCTCACTGATTTT 214
DB      121 CCCCTGTGTGGGCGTCTCATCAAAACCCAGCTGGGTGCTGGCCCAAGCTCACTGATTTT 180
QY      215 ACCCAATCTGAAGATGATCTGGAAATTTCAAGACAGAGTCAAGACGGTACTGAACA 274
DB      181 ACCCAATCTGAAGATGATCTGGAAATTTCAAGACAGAGTCAAGACGGTACTGAACA 240
QY      275 GACCAATTAACCCCATTCAGATGTCGCTCACTGGAACCTACATGATGAGGCCCCACAGGA 334
DB      241 GACCAATTAACCCCATTCAGATGTCGCTCACTGGAACCTACATGATGAGGCCCCACAGGA 300
QY      335 TGACCTCATGCTCATAGCTGGCTGAAGCTGTCATGCTCAATCCAAATGCCAGCCCT 394

```


Dh 301 TGACCTCATGCTCATCAAGCTGGCTAAGCTCCATCTCAATCCAAATCCAGCCCT 360
Qy 395 TCCCTGCGCCACCAATATGTCAGGCGAGGCACTGTCTGTACTCTCAGGTTGACTG 454
Db 361 TCCCTGCGCCACCAATATGTCAGGCGAGGCACTGTCTGTACTCTCAGGTTGACTG 420
Qy 455 GAGCGAAGAAAAGTGGCGGACACCTGACTTGGCGGAGAACTGGAGGCCCGCTGAT 514
Db 421 GAGCGAAGAAAAGTGGCGGACACCTGACTTGGCGGAGAACTGGAGGCCCGCTGAT 480
Qy 515 GTCTGATCGGAATGCCAAAAACAGAAACAGAAAGAACCAAGGAATTCCTATGCT 574
Db 481 GTCTGATCGGAATGCCAAAAACAGAAACAGAAAGAACCAAGGAATTCCTATGCT 540
Qy 575 GAAATTTGTGAATATTTAGCCGCAATTTTGGGAGAGTGGCCCTTCTACTGTCATCTG 634
Db 541 GAAATTTGTGAATATTTAGCCGCAATTTTGGGAGAGTGGCCCTTCTACTGTCATCTG 600
Qy 635 CAAAGCAAGCTCCAGGGAATCGAGTGGGCGACTTCATGAGGAGGGAACGTCGCACTGA 694
Db 601 CAAAGCAAGCTCCAGGGAATCGAGTGGGCGACTTCATGAGGAGGGAACGTCGCACTGA 660
Qy 695 CACCAATGTTACAAATATGTCATCTGATTTGAGAATGCTTAAGCAAGTGAAGACC 754
Db 661 CACCAATGTTACAAATATGTCATCTGATTTGAGAATGCTTAAGCAAGTGAAGACC 720
Qy 755 TACTTCTCCC 764
Db 721 TACTTCTCCC 730

RESULT 10
AAS44567
ID AAS44567 standard; cDNA; 708 BP.

AC AAS44567;

DT 18-DEC-2001 (first entry)

DE Human cDNA encoding protein AFP80526.

XX Human; sex: AFR; cytostatic; Marfan's syndrome; thrombocytopaenia;
XX leukaemia; porphyria; Gilles De la Tourette's syndrome; gene therapy;
XX squamous cell carcinoma; diabetes mellitus; Grave's disease;
XX colon cancer; Alzheimer's disease; epiphyseal dysplasia.

OS Homo sapiens.

PN W020016748-A2.

PD 13-SEP-2001.

PF 05-MAR-2001; 2001MO-US07192.

PR 03-MAR-2000; 2000US-187221P.

PA (ZYMO) ZYMOGENETICS INC.

PI Conklin DC, Presnell SR, Adler DA;

XX WPI; 2001-589943/66.

DR P-PSDB; AAU27666.

PT Novel AFP polypeptides and polynucleotides, useful for diagnostic and
PT therapeutic purposes, in cancer therapy and for screening modulator

PT compounds -

PS Claim 10; Page 192-193; 220pp; English.

CC The invention relates to novel human AFP proteins (not defined) and
CC the nucleic acids that encode them. AFP proteins are useful as standards
CC in assays of protein and protein inhibitors in both clinical and research
CC settings, as protein and amino acid supplements, including hydrolysates.

CC The nucleic acids are useful for radiation hybrid mapping. Secretory
CC fusion proteins of AFP are useful in cancer therapy, for enhancing in
CC vitro cytotoxicity, for enhancing in vivo killing of target tissues, and
CC for targeted cells or tissue inhibition or ablation. Anti-AFP antibodies
CC are useful for isolating target polypeptides by affinity purification, in
CC diagnostic assays for determining circulating or localised levels of
CC target polypeptides, for tissue typing, for cell sorting, for screening
CC expression libraries, for generating anti-idiotypic antibodies, and as
CC neutralising antibodies or as antagonists to block protein activity in
CC vitro and in vivo. AFP proteins and nucleic acids may be used to
CC diagnose or treat (e.g. by gene therapy) diseases associated with the
CC malfunction of the AFP e.g. AFP16824 and Marfan's syndrome,
CC AFP576853/AFP3158 and thrombocytopaenia, leukaemia, porphyria, Gilles De
CC La Tourette's syndrome, AFP552829 and squamous cell carcinoma, diabetes
CC mellitus, Grave's disease, AFP64311 and colon cancer, AFP253034 and
CC Alzheimer's disease and AFP68580 and epiphyseal dysplasia. Many more
CC examples of diseases are given in the specification. The present
CC sequence encodes an AFP of the invention.

XX Sequence 708 BP; 188 A; 185 C; 168 G; 167 T; 0 other;

Qy Query Match 81.6%; Score 706.4; DB 22; Length 708;

Db Best Local Similarity 99.9%; Pred. No. 3.2e-196;

Matches 707; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 42 ATGAATATGCTCTTCTATTGAGTGTCTCGTGGAGCAATTTCTTTCGTCATCTCT 101

Db 1 ATGAATATGCTCTTCTATTGAGTGTCTCGTGGAGCAATTTCTTTCGTCATCTCT 60

Qy 102 GTTCAGAAAAGAACCCCTGCTCCCTATTGTTGTATCTCAAGTCTCAATCCCTGT 161

Db 61 GTTCAGAAAAGAACCCCTGCTCCCTATTGTTGTATCTCAAGTCTCAATCCCTGT 120

Qy 162 GTGGGCGTCTCATCAACCCAGCTGGGTGGGCGCCAGCTCAGCTATTACCAAT 221

Db 121 GTGGGCGTCTCATCAACCCAGCTGGGTGGGCGCCAGCTCAGCTATTACCAAT 180

Qy 222 CTGAAGATGCTGGGAATTTCAAGACAGAGTCAAGAGCGTACTGAACAGCAAT 281

Db 181 CTGAAGATGCTGGGAATTTCAAGACAGAGTCAAGAGCGTACTGAACAGCAAT 240

Qy 282 AACCCATTCAGATCGTCCGCTACTGGAACATCAGTCAATAGGCCCCACAGATGACCT 341

Db 241 AACCCATTCAGATCGTCCGCTACTGGAACATCAGTCAATAGGCCCCACAGATGACCT 300

Qy 342 ATGCTCATCAAGTGGCTAAAGCTGCTCAATCCCAAGTCCAGGCCCTTCCCTC 401

Db 301 ATGCTCATCAAGTGGCTAAAGCTGCTCAATCCCAAGTCCAGGCCCTTCCCTC 360

Qy 402 GCCACCAATGTCAGGCGAGGCACTGTCTACTCTCAGTTTGGACTGAGCCAA 461

Db 361 GCCACCAATGTCAGGCGAGGCACTGTCTACTCTCAGTTTGGACTGAGCCAA 420

Qy 462 GAAACAGTGGCCGACACCTGACTTGGCGGAGAACTGGAAGCCCCCGTATGTTGAT 521

Db 421 GAAACAGTGGCCGACACCTGACTTGGCGGAGAACTGGAAGCCCCCGTATGTTGAT 480

Qy 522 CGAATATGCCAAAACAGAAAGGAAAGCAAGCAATTCCTATGTTGAATT 581

Db 481 CGAATATGCCAAAACAGAAAGGAAAGCAAGCAATTCCTATGTTGAATT 540

Qy 582 GTGAAGATTCAGCCGAATTTTGGGAGTGGCCGTTGCTACTGTCATCTGCAAGAC 641

Db 541 GTGAAGATTCAGCCGAATTTTGGGAGTGGCCGTTGCTACTGTCATCTGCAAGAC 600

Qy 642 AAGCTCAGGGAATGAGGTGGGCACTTCATGGAAGGAGACGTCGCAATCTTAACCAAT 701

Db 601 AAGCTCAGGGAATGAGGTGGGCACTTCATGGAAGGAGACGTCGCAATCTTAACCAAT 660

Qy 702 GTTTACAATATGATCTCTGATTTGGAACATCTGCTTAAGCAAGTGA 749

Db 661 GTTTACAATATGATCTCTGATTTGGAACATCTGCTTAAGCAAGTGA 708

RESULT 11
AB04647
ID AB04647 standard; cDNA; 721 BP.
XX
XX AB04647;
AC
XX
XX 21-FEB-2002 (first entry)
DT
XX
XX Coding sequence of MOL6a variant 13373750.
DE
XX
XX MOL; G-coupled protein-receptor; cardiomyopathy; atherosclerosis;
XX cell signal processing; metabolic disorder; diabetes; cancer;
XX neurodegenerative disorder; immune disorder; cardiac disorder;
XX lung disease; autoimmune disease; developmental disorder; antidiabetic;
XX Cystostatic; Neuroprotective; Antiatherosclerotic; Immunosuppressive;
XX Gene therapy; Vaccine; antiinflammatory; MOL6a; trypsin-like; ss.
XX
XX Unidentified.
OS
XX
XX Key Location/Qualifiers
FH 5..691
FT CDS
FT
FT variation /tag= a
FT /product= "MOL6a variant"
FT /replace(359,C)
XX /tag= b
XX
XX MO200181578-A2.
XX
XX 01-NOV-2001.
XX
XX 26-APR-2001; 2001MO-US13578.
XX
XX 26-APR-2000; 2000US-200158P.
XX 28-APR-2000; 2000US-200613P.
XX 28-APR-2000; 2000US-200780P.
XX 01-MAY-2000; 2000US-201006P.
XX 01-MAY-2000; 2000US-201007P.
XX 01-MAY-2000; 2000US-201236P.
XX 01-MAY-2000; 2000US-201238P.
XX 02-MAY-2000; 2000US-201186P.
XX 03-MAY-2000; 2000US-201474P.
XX 03-MAY-2000; 2000US-201508P.
XX 25-JUL-2000; 2000US-220591P.
XX 15-SEP-2000; 2000US-232678P.
XX 22-JAN-2001; 2001US-263217P.
XX 30-JAN-2001; 2001US-265160P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Verneer CAM, Fernandes ER, Gerlach V, Shinkets RA, Malysankar UM;
XX Boldeg FL, Zethusen BD, Spytek KA, Majumder K, Tchernyev VI;
XX Padigaru M, Patturajan M, Burgess CE, Gangoli EA, Smithson G;
XX Rastelli L, Macdougall JR, Taupier RJ, Grose WM, Szekeres ES;
XX Alsobrook JB;
XX
XX MPI; 2002-049278/06.
XX P-PSDB; AAM47674.
XX
XX Novel G-protein coupled receptor-related polypeptides and
XX polynucleotides for diagnosing, preventing and treating cardiomyopathy,
XX atherosclerosis, disorders related to cell signal processing and for
XX identifying modulators -
XX
XX Example 1; Page 216; 227pp; English.
XX
XX The present invention relates to novel G-coupled protein-receptor related
XX proteins and coding sequences (MOLX, where X is a number from 1 to 10,
XX AB04589-AB04603 and AAM47659-AAM47673). MOLX proteins and coding
XX sequences are useful for treating or preventing a MOLX-associated
XX disorder, such as cardiomyopathy, atherosclerosis, disorders related to
XX cell signal processing and metabolic pathway modulation, diabetes and
XX cancer. Additionally, MOLX proteins and coding sequences are useful for

CC preventing and treating a variety of disorders including metabolic
CC disorders, nutritional oedema, chronic and hereditary pancreatitis,
CC obesity, infectious disease, anorexia, neurodegenerative disorders,
CC Alzheimer's disease, Parkinson's disease, stroke, immune disorders,
CC haematopoietic disorders and various dyslipidaemias, metabolic syndrome X
CC and wasting disorders associated with chronic diseases and cancers,
CC cardiac disorders, hypertension, hypercalcaemia, cirrhosis, angiodysplasia
CC and wound healing, trauma, glomerulonephritis, hyper and hypothyroidism,
CC multiple sclerosis, lung diseases including asthma, Crohn's disease,
CC scleroderma, autoimmune diseases, developmental disorders and neural tube
CC defects. The present sequence is a coding sequence for a variant of
CC MOL6a. MOL6a is a trypsin-like protein.
XX
XX
SQ Sequence 721 BP; 193 A; 191 C; 169 G; 168 T; 0 other;
Query Match 81.3%; Score 703.8; DB 24; Length 721;
Best Local Similarity 99.6%; Pred. No. 1,9e-195;
Matches 716; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 39 ACCATGAATATGCTTCTATTGGGTGCTCGCTGGGACATTTTCTTGTGACTCA 98
DB 2 ACCATGAATATGCTTCTATTGGGTGCTCGCTGGGACATTTTCTTGTGACTCA 61
QY 99 TCTGTCAGAAAGAGACCCCTGCTCCCTATTGGTGTACCTCAAGTCTCAACCC 158
DB 62 TCTGTCAGAAAGAGACCCCTGCTCCCTATTGGTGTACCTCAAGTCTCAACCC 121
QY 159 TGTGTGGGCGTCTCTCAATCAACCCAGCTGGTGTCTGCCCCAGCTACTGCTATTACCA 218
DB 122 TGTGTGGGCGTCTCTCAATCAACCCAGCTGGTGTCTGCCCCAGCTACTGCTATTACCA 181
QY 219 AATCTGAAGTATGCTGTGGGAATTTCAAGAGAGAGTCAAGAGGTTCTGAACAGACA 278
DB 182 AATCTGAAGTATGCTGTGGGAATTTCAAGAGAGAGTCAAGAGGTTCTGAACAGACA 241
QY 279 ATTAACCCCATTCAGATCGTCCGCTACTGAACTACATGACCCGCCACAGATGAC 338
DB 242 ATTAACCCCATTCAGATCGTCCGCTACTGAACTACATGACCCGCCACAGATGAC 301
QY 339 CTCATGCTCATCAAGCTGGCTAAAGCTGCAATGCTCAATCCCAAGTCCAGCCCTTACC 398
DB 302 CTCATGCTCATCAAGCTGGCTAAAGCTGCAATGCTCAATCCCAAGTCCAGCCCTTACC 361
QY 399 CTCGCCACCAAGTCAAGGCGACGCTGCTGCTACTCTCAGATTGAGCTGAGC 458
DB 362 CTCGCCACCAAGTCAAGGCGACGCTGCTGCTACTCTCAGATTGAGCTGAGC 421
QY 459 CAAGAAACAGTGGCCGACACCTGACTTGGGAGAGAGCCCTGGAGGCCCCGATGCT 518
DB 422 CAAGAAACAGTGGCCGACACCTGACTTGGGAGAGAGCCCTGGAGGCCCCGATGCT 481
QY 519 GATCGAAGATGCC-AAAAACAGAACAGAAAGAAAGCCACAGAAATTCCTTATGTGAA 577
DB 482 GATCGAAGATGCC-AAAAACAGAACAGAAAGAAAGCCACAGAAATTCCTTATGTGAA 541
QY 578 ATTGTGAAGATATTCAGCCGAATTTTGGGAGGCTGCTGCTACTGATCTTGCAA 637
DB 542 ATTGTGAAGATATTCAGCCGAATTTTGGGAGGCTGCTGCTACTGATCTTGCAA 601
QY 638 AGACAAGCTCCAGGAATGAGGTGGGCACTTCATGAGAGGGAGCTGGCATCTACAC 697
DB 602 AGACAAGCTCCAGGAATGAGGTGGGCACTTCATGAGAGGGAGCTGGCATCTACAC 661
QY 698 CAATGTTTCAAAATATGATCTCTGATTGAGAACACTGCTTAAGAGACAGTACCTTA 756
DB 662 CAATGTTTCAAAATATGATCTCTGATTGAGAACACTGCTTAAGAGACAGTACCTTA 720
RESULT 12
ABK31786
ID ABK31786 standard; DNA; 798 BP.
XX
XX ABK31786;

XX Conklin DC;
 PI MPI; 2001-589946/66.
 DR Novel Ztryp3 polypeptides and polynucleotides useful in the treatment
 XX of asthma, vascular disorders including stroke, inflammation and
 PT testicular function
 PT
 XX
 PS Disclosure; Page 82; 82pp; English.
 XX
 CC The invention relates to an isolated human Ztryp3 polypeptide, a member
 CC of the serine protease family. Ztryp3 polypeptides and their associated
 CC polynucleotides are useful in diagnosis, therapy and industry and are
 CC used as targets for identifying modulators, preferably inhibitors of
 CC serine protease activity. The sequences are useful in the treatment of
 CC asthma, vascular function such as stroke, inflammation and testicular
 CC function (by modulating spermatogenesis). Ztryp3 proteins can be used for
 CC identifying peptide cleavage sites and for coupling amino and carboxy
 CC terminal tags. The polypeptides are also useful to teach analytical
 CC skills such as mass spectrometry, circular dichroism, X-ray
 CC crystallography and nuclear magnetic resonance spectroscopy. This
 CC sequence represents degenerate DNA encoding human Ztryp3.
 XX
 SQ Sequence 705 BP; 133 A; 68 C; 116 G; 102 T; 286 other;
 Query Match 59.1%; Score 511.4; DB 22; Length 705;
 Best Local Similarity 59.3%; Pred. No. 3.6e-139;
 Matches 418; Conservative 164; Mismatches 123; Indels 0; Gaps 0;
 QY 42 ATGAATATGCTCTTCTGAGTGGTCTGCTGAGGACATTTTCTGCTGATCATCT 101
 DB 1 ATGAATATGCTCTTCTGAGTGGTCTGAGGACATTTTCTGCTGATCATCT 60
 QY 102 GTTCAGAAAGAACCCCTGCTCTCTATTTGGTGTACCTCAAGTCTCAATCAACCCCTGT 161
 DB 61 GTTCAGAAAGAACCCCTGCTCTCTATTTGGTGTACCTCAAGTCTCAATCAACCCCTGT 120
 QY 162 GTGGGGGCTCTCTCAACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 221
 DB 121 GTGGGGGCTCTCTCAACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 QY 222 CTGAAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 281
 DB 181 YTNAAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
 QY 282 AACCCCATTCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 341
 DB 241 AAYCCNATCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 QY 342 ATGCTCATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 401
 DB 301 ATGCTCATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 QY 402 GCACACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 461
 DB 361 GCACACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 QY 462 GAAACAGTGGCGACACCTGACTGCGGACAGACCTGAGAGGCCCGCTGATGCTGAT 521
 DB 421 GAAACAGTGGCGACACCTGACTGCGGACAGACCTGAGAGGCCCGCTGATGCTGAT 480
 QY 522 CGAGATGCGCAAAAAACAGAACAGAAAAAGCAGACGAAATCTTATGTGTGAATTT 581
 DB 481 MNGAGATGTCARAAACAGAACAGAAAAAGCAGACGAAATCTTATGTGTGAATTT 540
 QY 582 GTGAAGATGTCAGCGCAATTTTGGGAGGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 641
 DB 541 GTGAAGATGTCAGCGCAATTTTGGGAGGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 QY 642 AAGCTTCAGAGGATCAGGTGGGCACTTCATGGAGGGGACCTCGGCACTTCAACCAAT 701
 DB 601 AARATTCAGAGGATCAGGTGGGCACTTCATGGAGGGGACCTCGGCACTTCAACCAAT 660

QY 702 GTTTCAAATATGATTCCTGATTCGAGACACTGCTAGACAG 746
 DB 661 GTTTCAAATATGATTCCTGATTCGAGACACTGCTAGACAG 705
 RESULT 14
 ID ABQ77083
 XX ABQ77083 standard; DNA; 705 BP.
 AC
 XX ABQ77083;
 DT 01-APR-2003 (first entry)
 XX
 DE Human serine protease Ztryp3 coding region.
 XX
 XX Human; serine protease; Ztryp3; gene; blood coagulation; fibrinolysis;
 KW complement activation; fertilization; hormone production; gene therapy;
 KW somatic cell; ds.
 XX
 OS Homo sapiens.
 XX
 XX US6468776-B1.
 PN 22-OCT-2002.
 PD
 XX 28-FEB-2001; 2001US-0796110.
 PF 03-MAR-2000; 2000US-186623P.
 PR
 XX (ZYMO) ZYMOGENETICS INC.
 PA
 XX Conklin DC;
 PI MPI; 2003-182091/18.
 DR P-PSDB; ABG73741.
 XX
 PT Novel serine protease nucleic acid molecule useful for producing serine
 PT protease polypeptide, and as educational tool in laboratory practicum
 PT kits for courses related to genetics and molecular biology
 PT
 XX
 PS Disclosure; Column 59-60; 32pp; English.
 XX
 CC This invention describes a novel human serine protease, Ztryp3. The
 CC serine protease family play a role in carefully controlled processes such
 CC as blood coagulation, fibrinolysis, complement activation, fertilization
 CC and hormone production. The products of the invention are useful for
 CC producing serine protease polypeptides; as probes or primers to clone 5'
 CC non-coding regions of Ztryp3 gene; as polymerase chain reaction primers
 CC and probes; as educational tool in laboratory practicum kits for courses
 CC related to genetics and molecular biology, protein chemistry, and
 CC antibody production and analysis; as standards or as unknowns for testing
 CC purposes; as an aid to teach a student how to prepare expression
 CC constructs for bacterial, viral, or mammalian expression, including
 CC fusion constructs, where Ztryp3 is the gene to be expressed; for
 CC determining the restriction endonuclease cleavage sites of the
 CC polynucleotides; for determining mRNA and DNA localization of Ztryp3
 CC polynucleotides in tissues (i.e., by northern and southern blotting as
 CC well as polymerase chain reaction); and for identifying related
 CC polynucleotides by nucleic acid hybridization. The products of the
 CC invention can also be used for determining mutations in the Ztryp3 gene,
 CC for determining Ztryp3 gene expression, for examining Ztryp3 gene
 CC structure, for detecting and localizing expression of Ztryp3 gene in a
 CC biological sample, in vivo diagnosis, in gene therapy, particularly
 CC in somatic cell gene therapy, and for producing transgenic mice. This
 CC sequence represents the degenerate coding region of the human serine
 CC protease Ztryp3 described in the method of the invention.
 XX
 SQ Sequence 705 BP; 133 A; 68 C; 116 G; 102 T; 286 other;
 Query Match 59.0%; Score 511.2; DB 25; Length 705;
 Best Local Similarity 59.3%; Pred. No. 4.1e-139;
 Matches 418; Conservative 163; Mismatches 124; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2003, 21:51:35 ; Search time 78 Seconds
(without alignments)
4900.485 Million cell updates/sec

Title: US-10-037-270-482

Perfect score: 866
Sequence: 1 ggcacttactccctgagctca.....tgggaatgagcatcactagt 866

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 113956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
4: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
5: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
6: /cgn2_6/prodata/2/ina/5B_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	866	100.0	866	4	US-09-620-312D-482
2	850.8	98.2	865	4	US-09-796-110-1
3	834	96.3	865	4	US-09-280-116-63
4	511.2	59.0	705	4	US-09-796-110-3
5	145	16.7	825	3	US-09-120-582-1
6	139.6	16.1	897	2	US-08-956-267A-1
7	66	7.6	1122	4	US-09-205-258-189
8	65.6	7.6	1146	4	US-09-205-258-247
9	65	7.5	452	4	US-09-280-116-64
10	64.8	7.5	1052	4	US-09-386-642-10
11	64.8	7.5	1166	3	US-08-944-483-7
12	64.8	7.5	1192	3	US-08-944-483-8
13	64.8	7.5	1314	3	US-09-025-059-2
14	63.4	7.3	833	2	US-08-790-137-2
15	63	7.3	732	1	US-08-361-395-2
16	60.6	7.0	725	3	US-09-197-801-12
17	60.6	7.0	725	3	US-09-551-028-12
18	60.6	6.4	969	3	US-09-502-600-30
19	55.2	6.4	986	2	US-08-557-146-1
20	55.2	6.4	986	2	US-09-154-344-1
21	55.2	6.4	1089	3	US-08-930-188-1
22	55.2	6.4	1089	3	US-08-930-188-3
23	55.2	6.4	1089	5	PCT-US96-04294-1
24	55.2	6.4	1089	5	PCT-US96-04294-3
25	54	6.2	1386	2	US-08-756-506-3
26	54	6.2	1387	6	5270178-1
27	54	6.2	1755	6	5225537-1

28	49.6	5.7	681	1	US-07-929-198-1	Sequence 1, Appl1
29	49.6	5.7	681	1	US-07-929-198-3	Sequence 3, Appl1
30	48	5.5	681	1	US-07-929-198-5	Sequence 5, Appl1
31	47.6	5.5	734	1	US-08-650-129-1	Sequence 1, Appl1
32	47.6	5.5	734	3	US-08-984-417-1	Sequence 1, Appl1
33	47.6	5.5	821	1	US-08-650-129-2	Sequence 2, Appl1
34	47.6	5.5	821	3	US-08-984-417-2	Sequence 2, Appl1
35	47.6	5.5	866	1	US-08-650-129-3	Sequence 3, Appl1
36	47.6	5.5	866	3	US-08-984-417-3	Sequence 3, Appl1
37	46.8	5.4	782	4	US-09-229-151C-9	Sequence 9, Appl1
38	46.8	5.4	833	4	US-09-229-151C-10	Sequence 10, Appl1
39	46.8	5.4	839	4	US-09-229-151C-8	Sequence 8, Appl1
40	46.8	5.4	884	2	US-08-851-974-2	Sequence 2, Appl1
41	46.8	5.4	884	2	US-09-213-390-2	Sequence 2, Appl1
42	46.8	5.4	2270	4	US-09-229-151C-13	Sequence 13, Appl1
43	46.8	5.4	2570	4	US-09-229-151C-14	Sequence 14, Appl1
44	46.4	5.4	693	4	US-09-402-515A-17	Sequence 17, Appl1
45	46.4	5.4	957	1	US-08-684-862-11	Sequence 11, Appl1

ALIGNMENTS

RESULT 1
US-09-620-312D-482

Sequence 482, Application US/09620312D

Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Zhang, Jie

APPLICANT: Ren, Feiyang

APPLICANT: Chen, Rui-hong

APPLICANT: Zhao, Qing A.

APPLICANT: Wehrman, Tom

APPLICANT: Xue, Aidong J.

APPLICANT: Yang, Yonghong

APPLICANT: Wang, Jian-Rui

APPLICANT: Zhou, Ping

APPLICANT: Ma, Yungting

APPLICANT: Wang, Dunrui

APPLICANT: Wang, Zhiwei

APPLICANT: John Tillinghast

APPLICANT: Drmanac, Radoje T.

TITLE OF INVENTION: No. 6569662el Nucleic Acids and

FILE REFERENCE: Polypeptides

CURRENT APPLICATION NUMBER: US/09/620,312D

PRIOR FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/552,317

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/488,725

NUMBER OF SEQ ID NOS: 1105

SOFTWARE: pt_FL_genes Version 1.0

SEQ ID NO 482

LENGTH: 866

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (42)..(749)

US-09-620-312D-482

Query Match

Best Local Similarity 100.0%; Score 866; DB 4; Length 866;

Best Local Similarity 100.0%; Pred. No. 7e-269;

Matches 866; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggcacttactccctgagctcaaggggagagagctgacacatgaaatgctttatt 60

Db 1 ggcacttactccctgagctcaaggggagagagctgacacatgaaatgctttatt 60

```

QY 61 TGGGTGTCCTCGCTGGAGCATTTTCTTGTGCTGATCTATCTGTTGAGAAAGAACCCCTG 120
DB 61 TGGGTGTCCTCGCTGGAGCATTTTCTTGTGCTGATCTATCTGTTGAGAAAGAACCCCTG 120
QY 121 CTCCTTATTTGGTGTACTCTCAAGTCTCACTTCAACCCCTGTGTGGGGTCTCTATCAAC 180
DB 121 CTCCTTATTTGGTGTACTCTCAAGTCTCACTTCAACCCCTGTGTGGGGTCTCTATCAAC 180
QY 181 CCAAGCTGGTGTGGGCCCCCAAGCTCACTGCTATTATTAACCAATTCGAAGTGTGGGAA 240
DB 181 CCAAGCTGGTGTGGGCCCCCAAGCTCACTGCTATTATTAACCAATTCGAAGTGTGGGAA 240
QY 241 ATTTCAAGACAGATGAGAGAGCGGTACTGAACAGCAATTAACCCCATTCAGATCTCC 300
DB 241 ATTTCAAGACAGATGAGAGAGCGGTACTGAACAGCAATTAACCCCATTCAGATCTCC 300
QY 301 GCTACTGGAATCTACATGATAGCGCCCAAGAGTACCTCATGCTCATCAAGTGGCTA 360
DB 301 GCTACTGGAATCTACATGATAGCGCCCAAGAGTACCTCATGCTCATCAAGTGGCTA 360
QY 361 AGCTGCGCATGCTCAATCCCAAGTCCAGGCCCTTCCCTGGCCAGCAATGTCAAGC 420
DB 361 AGCTGCGCATGCTCAATCCCAAGTCCAGGCCCTTCCCTGGCCAGCAATGTCAAGC 420
QY 421 CAGGCACTGTCTGTCTACTCTCAAGTGTGGAAGTGAAGCAAGAAACAGTGGCCGACCC 480
DB 421 CAGGCACTGTCTGTCTACTCTCAAGTGTGGAAGTGAAGCAAGAAACAGTGGCCGACCC 480
QY 481 CTGACTTGGCGGAGAACTGTGAAGGCCCGGTGATGTGTGATTCGAATGCCAAAAACAG 540
DB 481 CTGACTTGGCGGAGAACTGTGAAGGCCCGGTGATGTGTGATTCGAATGCCAAAAACAG 540
QY 541 AACAGGAAAAAGCCACAGGAATTCCTTATGTGTGAATTTGTGAAGTATTCAGCCGAA 600
DB 541 AACAGGAAAAAGCCACAGGAATTCCTTATGTGTGAATTTGTGAAGTATTCAGCCGAA 600
QY 601 TTTTGGGAGGTGGCCGTTGCTACTGTCACTGTGCAAGCAAGCAAGCTTCAGGGAATCGAG 660
DB 601 TTTTGGGAGGTGGCCGTTGCTACTGTCACTGTGCAAGCAAGCAAGCTTCAGGGAATCGAG 660
QY 661 TGGGGCACTTCATGGGAGGAGCGTGGCATCTACCAATGTTTACAAATATGTATCTT 720
DB 661 TGGGGCACTTCATGGGAGGAGCGTGGCATCTACCAATGTTTACAAATATGTATCTT 720
QY 721 GGAATGGAACACTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB 721 GGAATGGAACACTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 781 TCTGCCATGAGACTATACAGCAGATATTTTCCCTCTATTCAAATAATTCCTCAATG 840
DB 781 TCTGCCATGAGACTATACAGCAGATATTTTCCCTCTATTCAAATAATTCCTCAATG 840
QY 841 AAAATTTGGGAATGTAGCATACTAGT 866
DB 841 AAAATTTGGGAATGTAGCATACTAGT 866

```

RESULT 2
US-09-796-110-1
; Sequence 1, Application US/09796110
; Patent No. 6468776
; GENERAL INFORMATION:
; APPLICANT: Konklin, Darrell C.
; TITLE OF INVENTION: Human Serine Protease
; FILE REFERENCE: 00-16
; CURRENT APPLICATION NUMBER: US/09/796.110
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FaastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 865
; TYPE: DNA
; ORGANISM: Homo sapiens

FEATURE:
; NAME/KEY: CDS
; LOCATION: (42) ... (746)
US-09-796-110-1

Query Match 98.2%; Score 850.8; DB 4; Length 865;
Best Local Similarity 99.7%; Pred. No. 5,5e-264;
Matches 863; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```

QY 1 GGCACCTACTCCCTGAGCTTAAGGGGGAAGAGAGCTGATACCAATGAAATATGCTTATTT 60
DB 1 GGCACCTACTCCCTGAGCTTAAGGGGGAAGAGAGCTGATACCAATGAAATATGCTTATTT 60
QY 61 TGGGTGTCCTCGCTGGAGCATTTTCTTGTGCTGATCTATCTGTTGAGAAAGAACCCCTG 120
DB 61 TGGGTGTCCTCGCTGGAGCATTTTCTTGTGCTGATCTATCTGTTGAGAAAGAACCCCTG 120
QY 121 CTCCTTATTTGGTGTACTCTCAAGTCTCACTTCAACCCCTGTGTGGGGTCTCTATCAAC 180
DB 121 CTCCTTATTTGGTGTACTCTCAAGTCTCACTTCAACCCCTGTGTGGGGTCTCTATCAAC 180
QY 181 CCAAGCTGGTGTGGGCCCCCAAGCTCACTGCTATTATTAACCAATTCGAAGTGTGGGAA 240
DB 181 CCAAGCTGGTGTGGGCCCCCAAGCTCACTGCTATTATTAACCAATTCGAAGTGTGGGAA 240
QY 241 ATTTCAAGACAGATGAGAGAGCGGTACTGAACAGCAATTAACCCCATTCAGATCTCC 300
DB 241 ATTTCAAGACAGATGAGAGAGCGGTACTGAACAGCAATTAACCCCATTCAGATCTCC 300
QY 301 GCTACTGGAATCTACATGATAGCGCCCAAGAGTACCTCATGCTCATCAAGTGGCTA 360
DB 301 GCTACTGGAATCTACATGATAGCGCCCAAGAGTACCTCATGCTCATCAAGTGGCTA 360
QY 361 AGCTGCGCATGCTCAATCCCAAGTCCAGGCCCTTCCCTGGCCAGCAATGTCAAGC 420
DB 361 AGCTGCGCATGCTCAATCCCAAGTCCAGGCCCTTCCCTGGCCAGCAATGTCAAGC 420
QY 421 CAGGCACTGTCTGTCTACTCTCAAGTGTGGAAGTGAAGCAAGAAACAGTGGCCGACCC 480
DB 421 CAGGCACTGTCTGTCTACTCTCAAGTGTGGAAGTGAAGCAAGAAACAGTGGCCGACCC 480
QY 481 CTGACTTGGCGGAGAACTGTGAAGGCCCGGTGATGTGTGATTCGAATGCCAAAAACAG 540
DB 481 CTGACTTGGCGGAGAACTGTGAAGGCCCGGTGATGTGTGATTCGAATGCCAAAAACAG 540
QY 541 AACAGGAAAAAGCCACAGGAATTCCTTATGTGTGAATTTGTGAAGTATTCAGCCGAA 600
DB 541 AACAGGAAAAAGCCACAGGAATTCCTTATGTGTGAATTTGTGAAGTATTCAGCCGAA 600
QY 601 TTTTGGGAGGTGGCCGTTGCTACTGTCACTGTGCAAGCAAGCAAGCTTCAGGGAATCGAG 660
DB 601 TTTTGGGAGGTGGCCGTTGCTACTGTCACTGTGCAAGCAAGCAAGCTTCAGGGAATCGAG 660
QY 661 TGGGGCACTTCATGGGAGGAGCGTGGCATCTACCAATGTTTACAAATATGTATCTT 720
DB 661 TGGGGCACTTCATGGGAGGAGCGTGGCATCTACCAATGTTTACAAATATGTATCTT 720
QY 721 GGAATGGAACACTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB 721 GGAATGGAACACTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 781 TCTGCCATGAGACTATACAGCAGATATTTTCCCTCTATTCAAATAATTCCTCAATG 840
DB 781 TCTGCCATGAGACTATACAGCAGATATTTTCCCTCTATTCAAATAATTCCTCAATG 840
QY 841 AAAATTTGGGAATGTAGCATACTAGT 866
DB 841 AAAATTTGGGAATGTAGCATACTAGT 866

```

RESULT 3
US-09-280-116-63/c
; Sequence 63, Application US/09280116A

```

; Patent No. 6331427
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
; FILE REFERENCE: 5800-24, 035800/176965
; CURRENT APPLICATION NUMBER: US/09/280,116A
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 63
; LENGTH: 865
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: trypsin-like serine proteases
US-09-280-116-63

Query Match          96.3%; Score 834; DB 4; Length 865;
Best Local Similarity 98.7%; Pred. No. 1,4e-258;
Matches 851; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

QY 1 GGCACCTTACCTCCCTGAGCTAAAGGGGAGAGAGCTGATCACCATGAATATGCTCTTATT 60
DB 865 GGCACCTTACCTCCCTGAGCTAAAGGGGAGAGAGCTGATCACCATGAATATGCTCTTATT 806
QY 61 TGGGTGCTCCTCGCTGGAGACATTTTCTTTGCTGACTCATCTGTTCAAGAAAGAACCCCTG 120
DB 805 TGGGTGCTCCTCGCTGGAGACATTTTCTTTGCTGACTCATCTGTTCAAGAAAGAACCCCTG 746
QY 121 CTCCTTAATTTGATGTAAGCTCAAGCTTCAACCCCTGTGTGGGCGTCTCATCAAC 180
DB 745 CTCCTTAATTTGATGTAAGCTCAAGCTTCAACCCCTGTGTGGGCGTCTCATCAAC 686
QY 181 CCAAGCTGGGTGCTGGCCCGGCTCAGTCTGCTATTACCAATCTGAAAGTATGCTGGAA 240
DB 685 CCAAGCTGGGTGCTGGCCCGGCTCAGTCTGCTATTACCAATCTGAAAGTATGCTGGAA 626
QY 241 ATTTCAAGAGCAGAGTCAGAGACGCTACTGAACAGACAAATTAACCCCATTCAGATGCTC 300
DB 625 ATTTCAAGAGCAGAGTCAGAGACGCTACTGAACAGACAAATTAACCCCATTCAGATGCTC 566
QY 301 GCTACTGGAACCTACAGTCATAGGCGCCCAAGAGATGACCTCATGCTCATCAAGCTGACTA 360
DB 565 GCTACTGGAACCTACAGTCATAGGCGCCCAAGAGATGACCTCATGCTCATCAAGCTGACTA 506
QY 361 AGCTTCGCATGCTCAATCCCAAGTCCAGCCCTT-CCCTCGCCACACCAATGTCAGG 419
DB 505 AGCTTCGCATGCTCAATCCCAAGTCCAGCCCTT-CCCTCGCCACACCAATGTCAGG 446
QY 420 CCAAGGCACTGCTGTAAGTCTCAGGTTTGAAGTGAAGCAAGAAAGAGTGGCCGACAC 479
DB 445 CCAAGGCACTGCTGTAAGTCTCAGGTTTGAAGTGAAGCAAGAAAGAGTGGCCGACAC 386
QY 480 CCTGACTTGGGAGAGACCTGAGAGCCCGCTGATGCTGATGAGAAATGCCAAAAACA 539
DB 385 CCTGACTTGGGAGAGACCTGAGAGCCCGCTGATGCTGATGAGAAATGCCAAAAACA 326
QY 540 GAAACAAGAAAAAGCCACAGGAATTCCTTAATGTGTGAATTTGTGAAGATTCACCCGA 599
DB 325 GAAACAAGAAAAAGCCACAGGAATTCCTTAATGTGTGAATTTGTGAAGATTCACCCGA 266
QY 600 ATTTTGGGGAGAGTGGCCGCTGCTACTGTCATCTGAAAGAAAGCAAGCTCCAGGAAATCGAG 659
DB 265 ATTTTGGGGAGAGTGGCCGCTGCTACTGTCATCTGAAAGAAAGCAAGCTCCAGGAAATCGAG 206
QY 660 GTGGGGCACTTCAAGGAGAGGAGCTGCGCATCTACCAATGTTTAAATATGATCC 719
DB 205 GTGGGGCACTTCAAGGAGAGGAGCTGCGCATCTACCAATGTTTAAATATGATCC 146
QY 720 TGGATTGAGAACCTGCTAAAGAACAGTGAACCTTACTTCTCTGCTGATTCACCTGG 779
DB 145 TGGATTGAGAACCTGCTAAAGAACAGTGAACCTTACTTCTCTGCTGATTCACCTGG 86
```

```

QY 780 CTCCTGCCAGACTATACAGCAGATATTTTCCCTCTATTCATAAATAATCTCCAAAT 839
DB 85 CTCCTGCCAGACTATACAGCAGATATTTTCCCTCTATTCATAAATAATCTCCAAAT 26
QY 840 GAAATTTGGGAATGTAGCATA 861
DB 25 GAAATTTGGGAATGTAGCATA 4

RESULT 4
US-09-796-110-3
; Sequence 3, Application US/09796110
; Patent No. 6468776
; GENERAL INFORMATION:
; APPLICANT: Konklin, Darrell C.
; TITLE OF INVENTION: Human Serine Protease
; FILE REFERENCE: 00-16
; CURRENT APPLICATION NUMBER: US/09/796,110
; CURRENT FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 705
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This degenerate nucleotide sequence encodes the
; OTHER INFORMATION: amino acid sequence of SEQ ID NO12.
; NAME/KEY: misc_feature
; LOCATION: (1)...(705)
; OTHER INFORMATION: n = A,T,C or G
US-09-796-110-3

Query Match          59.0%; Score 511.2; DB 4; Length 705;
Best Local Similarity 59.3%; Pred. No. 1.2e-154;
Matches 418; Conservative 163; Mismatches 124; Indels 0; Gaps 0;

QY 42 ATGAATATGTCCTTCATTTGGGTGTCCTCGCTGGGACATTTTCTTGTGCTGATCT 101
DB 1 ATGAATATGTCCTTCATTTGGGTGTCCTCGCTGGGACATTTTCTTGTGTCGATGTC 60
QY 102 GTTCAGAAAGAGACCCCTCTCTCTATTTGGTACTCAAGTCTCAAGTCAACCCCTGT 161
DB 61 GTTCAGAAAGAGAGACCCCTCTCTCTATTTGGTACTCAAGTCTCAAGTCAACCCCTGT 120
QY 162 GTGGGCGTCTCATCAACCCAGCTGGGTGCTGGCCCAAGCTCAGTCTATTTACCAAT 221
DB 121 GTGGGCGTCTCATCAACCCAGCTGGGTGCTGGCCCAAGCTCAGTCTATTTACCAAT 180
QY 222 CTGAAGATGATCTGGGAAATTTCAAGACAGTCAAGAGAGGTCTGAACAGCAATT 281
DB 181 CTGAAGATGATCTGGGAAATTTCAAGACAGTCAAGAGAGGTCTGAACAGCAATT 240
QY 282 AACCCCATTCAGATGCTCGCTACTAGTGAAGTCAAGTCAAGAGAGGTCTGAAGTCACT 341
DB 241 AACCCCATTCAGATGCTCGCTACTAGTGAAGTCAAGTCAAGAGAGGTCTGAAGTCACT 300
QY 342 ATGCTCATCAAGCTGCTAAGCCTGCTGATGCTCAATCCCAAGTCCAGGCCCTTCCCTC 401
DB 301 ATGCTCATCAAGCTGCTAAGCCTGCTGATGCTCAATCCCAAGTCCAGGCCCTTCCCTC 360
QY 402 GCCACCAACCAAGTCAAGGACAGGCACTGCTCTACTCTCAGGTTTGAAGTGAAGCCAA 461
DB 361 GCCACCAACCAAGTCAAGGACAGGCACTGCTCTACTCTCAGGTTTGAAGTGAAGCCAA 420
QY 462 GAAACAAGTGGCGACACCTGACTTGGCGGAGAACTGAGAGGCCCGCTGATGTCAT 521
DB 421 GAAACAAGTGGCGACACCTGACTTGGCGGAGAACTGAGAGGCCCGCTGATGTCAT 480
QY 522 CGAAGATGCAAAAAACAGAACAGAAAAAGCCACAGGAATTCCTTATGTGTGAATTT 581
DB 481 MNGARTGVCABAABACNGARCARGNARBMNCAYMGNAAAYMSVNTNTGTYTNAARTTY 540
```

QY 582 TGGAAAGATTAGGCGAATTTTGGGAGAGGGCCGTTGCTACTGCAATCTGCAAGAAGAC 6411

Db 541 GTTAAAGTNTTWTSMNGAHTHTTYYGGNGARGTNGCAGTNGCAACGTNATHTHTYAAAGAY 6000

QY 642 AAGCTCCAGGAATGAGGTGGGGGCCTTCATGGAGGGGAGCGTCGGCATCTCAACCAAT 701

Db 601 AARVTCARCGAATHGARGTNGNCAYTTATGGGNGGNGAYGTNGGNAHTHTYACCAAY 6600

QY 702 GTTTACAATATGTATCCTGAGTTGGAACACATCGCTAAGGACAAG 746

Db 661 GTNTTAAARTAYGTNMSNTGATHGARRAAYACGCAAAAGATYAR 705

RESULT 5

US-09-120-582-1
 : Sequence 1, Application US/09120582
 : Patent No. 6087558
 : GENERAL INFORMATION:
 : APPLICANT: Howard, John A.
 : APPLICANT: Hood, Elizabeth
 : TITLE OF INVENTION: COMMERCIAL PRODUCTION OF PROTEASES IN PLANTS
 : FILE REFERENCE: 10014
 : CURRENT APPLICATION NUMBER: US/09/120,582
 : CURRENT FILING DATE: 1998-07-22
 : NUMBER OF SEQ ID NOS: 2
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 1
 : LENGTH: 825
 : TYPE: DNA
 : ORGANISM: Bos taurus
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: (44)..(730)
 US-09-120-582-1

Query Match	16.7%	Score	145	DB 3	Length	825			
Best Local Similarity	52.5%	Pred. NO.	1.3e-36						
Matches	347	Conservative	0	Mismatches	305	Indels	9	Gaps	1

Qy 111 GAAGACCCGTGCTCCCTATTGGTGTGTAACCTCAAGTCTCATTCTCAACCCCTGTGTGGGCGTC 170

Dp 86 GCAAATACTGTCTCCCTACCAAGTGTCCGTGAACCTGTGGGTACCACTTTGGGGGGCTCC 145

Qy 171 CTCATCAAAACCAAGCTGGGTGTGGGCCCACTCACTGCTATTATACAAATCTGAAAGTG 230

Dp 146 CTCATCAACAAGCAAGTGGGTGTGTCTGGGGCTCACTGGTCTACAAGTCCGAATCCAAAGTG 205

Qy 231 ATGTGGGAAATTTTCAAGACAGAGTCAGAGA CCGTACTGAACAGACAATTTACCCCAT 290

Dp 206 CGTCTGGAGAGAGACAACTTTATGTCTGTGAGGGCAATGAGCAATTCATCAGCGATCC 265

Qy 291 CAGATGTCCTGCTACTGGAACCTACAGTCATGAGGCCCAAGAGTACCTCATGCTATC 350

Dp 266 AAGAGTATGTCATCTCCACGTACAACTCAAAACCTTTAAACAAGCATCATGTGATTT 325

Qy 351 AAGCTGGCTAAGCCCTGCCATGCTCAATCCCAAGTCCACCCCTTCCCTGGCACACAC 410

Dp 326 AAACCTGAATCAGCTGCACAGTCTCAACAGCCGAGTAGCCTCTATCTCTCTGCCAATCC 385

Qy 411 AATGTCAAGCCACAGGACCTGTCTGTCTACTCTCAGTGTGGAATGAGCCACAAGAAAAGT 470

Dp 386 TGTGCTCTGTGCTGGCACCCAGTGTCTCATCTCTGGGTGGGGCAACAACAAAGAGCATGGG 445

Qy 471 GGCGCACACCCCTGACTTGGGGCAGAAACCTGAGAGCCCCCTGATGTCTGATGAGAAATGC 530

Dp 446 ACCACACTACCCCTGAGTGTCTGAAAGTGTCTGAAAGGCTCCCATCTTATCAACAACACTTTGC 505

Qy 531 CAAAAAACAAGAACAGAAAAAAGCCACAGAGAAATTCCTTATGTGTGAAATTTTGGAAAGTA 590

Dp 506 AAAAGTGTCTACCCAGGGCCAGATCAACAGCAATGTTCTGTGCGGGGTACCTCGAGGGG 565

Qy 591 TTCAGCCGAATTTTGGGAGAGGTGGCCGTTGCTACTGTCACTTGCACAAAGCAAGCTTCAG 650

[illegible]

RESULT 6

```

US-08-956-267A-1
; Sequence 1, Application US/08956267A
; Patent No. 5945328
;
; GENERAL INFORMATION:
;
; APPLICANT: WOLDIKE, Helle Fabricius
; APPLICANT: KJELDSEN, Thomas Borglum
; TITLE OF INVENTION: A Process for Producing Trypsin
; TITLE OF INVENTION: (Trypsinogen)
; NUMBER OF SEQUENCES: 6
;
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 59453280 No. 5945328disk of No. 5945328th America, Inc.
; STREET: 405 Lexington Avenue
;
; CITY: New York
; STATE: NY
;
; COUNTRY: USA
; ZIP: 10174
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,267A
; FILING DATE: 22-OCT-1997
;
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rozek, Carol
; REGISTRATION NUMBER: 36,993
; REFERENCE/DOCKET NUMBER: 4500.204-US
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 897 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
;
; FEATURE:
;
; NAME/KEY: CDS
; LOCATION: 4..744
;
US-08-956-267A-1

```

[illegible]

Db	219	GCCTCTGGAGAAACAAACATGACGTCCTTGAAGGCGAATAGCAATTCATCATGCGC	278
Qy	290	TCAGATTCGTCCTACTGGAACCTACAGTCANAGCCGCCACAGATGACTCATGCTCAT	349
Db	279	CAAGATCATACCAACCCCAATTTCAATGAAATACCTTAGTAACGACATCATGTGAT	338
Qy	350	CAAGCTGGCTAAGCCTTGCCTCATGTCTCAATCCCAAAATCCAGCCCTTCCCTCGCACAC	409
Db	339	TAACTGAGCTCACTCGCCACTCTCAAGCTGAGTACCACTGTCTCACTGCCAAGATC	398
Qy	410	CAATGTCAGGCGCAGGACGTCTGTCTACTCTCAGATTGGACTGAGGCCAAGAAAACAG	469
Db	399	TTGTGACAGCTGCTGTAACCGAAGTGTCTATCTTGGCTGGGGCAACACAAAGCAGTGG	458
Qy	470	TGCGCGAACACCTGACTTGCAGGACGAACCTGAGGCCGCCGCTGATGTCTGATCGAGATG	529
Db	459	CTCCAGCTACCTTTCGCTCTCGCAATAGCTGAAAGGCCCCCGCTCAAGTGAACATTTTG	518
Qy	530	CCAAAAACAGAACAGAAAAAACCAACAGAAATTCCTTAATGTGTGAATTTGTAAAGT	589
Db	519	CAAGAATTCCTACCCCGCCAGATCAACCGAAACATGATCTGTGTGGCTTCCGAGGG	578
Qy	590	ATTCAAGCCGAATTTTGGGGAAGGTGGCGGTGCTACTCATCTGCAAGCAAGCACTCCA	649
Db	579	TGTTAAGATTTCTTGCCAGGAAACTCTGTGTGCCCCGCGGTGTGCAATGACAGCTTCA	638
Qy	650	GGGAATCAGAGTGGGGGCACTT-----CATGGAGAGGACGTGGCACTTACACAA	700
Db	639	GGGTATTTGTCTCTTGGGGCTATGTGCTGCGCCCAAGAAAAACAAGCTGGGTCTACAA	698
Qy	701	TGTTTAAATATGATCTCTGATTTGAGAACCTGCTTAAGACAAAGTAGACCTTACTTC	760
Db	699	GGTCTCAACTATGTGACTGATTCAAGACATCGCTGCCAATTAAGATTTCAIT	758
Qy	761	TC 762	
Db	759	TC 760	

RESULT 7
 US-09-205-258-189
 Sequence 189, Application US/09205258
 Patent No. 6525174
 GENERAL INFORMATION:
 APPLICANT: Young et al.
 TITLE OF INVENTION: 207 Human Secreted Proteins
 FILE REFERENCE: P2007P1
 CURRENT APPLICATION NUMBER: US/09/205,258
 CURRENT FILING DATE: 1998-12-04
 EARLIER APPLICATION NUMBER: PCT/US98/11422
 EARLIER FILING DATE: 1998-06-04
 EARLIER APPLICATION NUMBER: 60/048,885
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/049,375
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/048,881
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/048,880
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/048,896
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/049,020
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/048,876
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/048,895
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/048,884
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/048,894
 EARLIER FILING DATE: 1997-06-06
 EARLIER APPLICATION NUMBER: 60/048,971

1	EARLIER	FILING DATE:	1997-06-06
2	EARLIER	APPLICATION NUMBER:	60/048,964
3	EARLIER	FILING DATE:	1997-06-06
4	EARLIER	APPLICATION NUMBER:	60/048,882
5	EARLIER	FILING DATE:	1997-06-06
6	EARLIER	APPLICATION NUMBER:	60/048,899
7	EARLIER	FILING DATE:	1997-06-06
8	EARLIER	APPLICATION NUMBER:	60/048,893
9	EARLIER	FILING DATE:	1997-06-06
10	EARLIER	APPLICATION NUMBER:	60/048,900
11	EARLIER	FILING DATE:	1997-06-06
12	EARLIER	APPLICATION NUMBER:	60/048,901
13	EARLIER	FILING DATE:	1997-06-06
14	EARLIER	APPLICATION NUMBER:	60/048,892
15	EARLIER	FILING DATE:	1997-06-06
16	EARLIER	APPLICATION NUMBER:	60/048,915
17	EARLIER	FILING DATE:	1997-06-06
18	EARLIER	APPLICATION NUMBER:	60/049,019
19	EARLIER	FILING DATE:	1997-06-06
20	EARLIER	APPLICATION NUMBER:	60/048,970
21	EARLIER	FILING DATE:	1997-06-06
22	EARLIER	APPLICATION NUMBER:	60/048,972
23	EARLIER	FILING DATE:	1997-06-06
24	EARLIER	APPLICATION NUMBER:	60/048,916
25	EARLIER	FILING DATE:	1997-06-06
26	EARLIER	APPLICATION NUMBER:	60/049,373
27	EARLIER	FILING DATE:	1997-06-06
28	EARLIER	APPLICATION NUMBER:	60/048,875
29	EARLIER	FILING DATE:	1997-06-06
30	EARLIER	APPLICATION NUMBER:	60/049,374
31	EARLIER	FILING DATE:	1997-06-06
32	EARLIER	APPLICATION NUMBER:	60/048,917
33	EARLIER	FILING DATE:	1997-06-06
34	EARLIER	APPLICATION NUMBER:	60/048,949
35	EARLIER	FILING DATE:	1997-06-06
36	EARLIER	APPLICATION NUMBER:	60/048,974
37	EARLIER	FILING DATE:	1997-06-06
38	EARLIER	APPLICATION NUMBER:	60/048,883
39	EARLIER	FILING DATE:	1997-06-06
40	EARLIER	APPLICATION NUMBER:	60/048,897
41	EARLIER	FILING DATE:	1997-06-06
42	EARLIER	APPLICATION NUMBER:	60/048,898
43	EARLIER	FILING DATE:	1997-06-06
44	EARLIER	APPLICATION NUMBER:	60/048,962
45	EARLIER	FILING DATE:	1997-06-06
46	EARLIER	APPLICATION NUMBER:	60/048,963
47	EARLIER	FILING DATE:	1997-06-06
48	EARLIER	APPLICATION NUMBER:	60/048,877
49	EARLIER	FILING DATE:	1997-06-06
50	EARLIER	APPLICATION NUMBER:	60/048,878
51	EARLIER	FILING DATE:	1997-06-06
52	EARLIER	APPLICATION NUMBER:	60/070,923
53	EARLIER	FILING DATE:	1997-12-18
54	EARLIER	APPLICATION NUMBER:	60/092,921
55	EARLIER	FILING DATE:	1998-07-15
56	EARLIER	APPLICATION NUMBER:	60/094,657
57	EARLIER	FILING DATE:	1998-07-30
58	NUMBER OF SEQ ID NOS:	1227	
59	SOFTWARE:	PatentIn Ver. 2.0	
60	SEQ ID NO 189		
61	LENGTH:	1292	
62	TYPE:	DNA	
63	ORGANISM:	Homo sapiens	
64	US-09-205-258-189		

Query Match Score 66; DB 4; Length 1292;
 Best Local Similarity 47.1%; Pred. No. 4.9e-11;
 Matches 299; Conservative 1; Mismatches 301; Indels 24; Gaps 2;

QY 216 CCAATCTGAAGTATGCTGGGAAATTTCAAGACAGAGTCAAGACGGTACTGAACAG 275
DB 402 CCCCGCTACATAGTTCACTGGGGCAGACAACTCCAGAGAGAGGGCTGTGACAG 461
QY 276 ACATTAATACCCATTCAGTGTCCGCTACTGGAACATACATCTACCGCCCAAGAT 335
DB 462 ACCCGAGACGACTGATGCTCTTCCCAACCCCGCTTCAACAAACCTCCCAACAA 521
QY 336 GACC-----TCATGCTCATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 383
DB 522 GACACCGGCAATGATCATGCTGTGTGAATGATGATGATGATGATGATGATGATGAT 581
QY 384 GTCCAGCCCTTCCCTCTGCGCAGCAACATGTCAGCGCAGGCACTGTCTGTACTCA 443
DB 582 GTGCGACCCCTCAGCT 641
QY 444 GGTTCGACTGAGCAGCAAGAAACAGTGGCCGACACCTGATCTGCGGCAAGACTGAG 503
DB 642 GGGTGGGAGCAGCTGACGCTCCAGCTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCT 701
QY 504 GCGCCCGTATGCTGATTCGGAATGCAAAACAGAAACAGAAACAGAAACAGAAACAG 563
DB 702 ATCAGCATCATGATGACACGAACTGTGAGAGCCCTTACCCCGCAATCAGAGAAC 761
QY 564 TCCTTATGTGTAATTTGTAAGTATTCAGCCGAAATTTTGGGAGGTGGCGTGTCT 623
DB 762 ATGGTGTGTGCGACGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 821
QY 624 ACTGTCATCTGCAAGACAGAGCTCCAGGAGT-----CGAGTGGGCACTTC 671
DB 822 CCTGTGCTGTAAACGATCTCTTCAAGGCAATATCTCTGCGGCGAGATCCGTGCG 881
QY 672 ATGGAGAGGAGCGTGGCATCTACACCAATGTTTAAATATGATCTGATGAGAAC 731
DB 882 ATCAGCGGAAAGCTGTGTCTTACAGAAAGTCTGCAATATGTGACTGATCAGAG 941
QY 732 ACTGCTAAGACAACT 747
DB 942 ACGATGAAGAACT 957

RESULT 8
US-09-205-258-247
Sequence 247, Application US/09205258
Patent No. 6525174
GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: P2007P1
CURRENT APPLICATION NUMBER: US/09/205, 258
CURRENT FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: PCT/US98/11422
EARLIER FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: 60/048, 885
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049, 375
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 881
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 880
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 896
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049, 020
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 876
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 884
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 894
EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048, 971
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 882
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 899
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049, 019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049, 373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049, 374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070, 923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092, 921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094, 657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 247
LENGTH: 1146
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (20)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (35)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:

NAME/KEY: SITE
 LOCATION: (36)
 OTHER INFORMATION: n equals a,t,g, or c
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (37)
 OTHER INFORMATION: n equals a,t,g, or c
 US-09-205-258-247

Query Match 7.6%; Score 65.6; DB 4; Length 1146;
 Best Local Similarity 46.9%; Pred. No. 6.1e-11;
 Matches 289; Conservative 2; Mismatches 301; Indels 24; Gaps 2;

QY 156 CCCTGTGTGGGCGCTCTCATCAAAACCCAGCTGGGTGGTGGCCCACTCACTCTATTAA 215
 DB 196 CTCTGTGGGGCGAGCGTCAATCGCCCAAGATGGCTCTGTACAGACAGCCCACTCCCAAG 255
 QY 216 CCAATCTGAAAGTATGCTGGGAAATTTCAAGACAGAGTCAAGACGGTACTGAACAG 275
 DB 256 CCCCCTAATAGTTACCTGGGGGACACAACTCCAGAGAGGAGGGCTGTAGCAG 315
 QY 276 ACAATTAACCCCTTTCAGATCCCTCCCTAATGGAATCAAGTATGAGCCGCCCAAGAT 335
 DB 316 ACCCGACAGCCACTGATCTTCCCAACCCGCTTCAACAAAGCCCTCCCAACAA 375
 QY 336 GACC-----TCATGCTCATCAAGCTGAGCTGAGCTGCAATGCTCAATCCCAAA 383
 DB 376 GACCAACCGCAATGATCATGCTGTGAGATGGATGGCATCCCACTTCCATCACTGGGCT 435
 QY 384 GTCCAGACCCCTTCCCTCCGACACCAATGTGAGGCGAGCACTGTCTGTACTCTCA 443
 DB 436 GTCCGACCCCTCACTCTCTCCAGCTGTGTACTGTGTGAGCAAGCTGTCTCATTTCC 495
 QY 444 GCTTTGACTGTGAGCCCAAGAAACAGTGGCCGACACCTGTACTGTGGGCAAACTGGAG 503
 DB 496 GGTGGGGGAGAGAGTCCAGCCCACTTACGCTGTCTCACTTGTGATGCGCAAC 555
 QY 504 GCCCCGCTGATGCTGATGAGATGCAAAACAGAAACAAAGAAAGCCACAGGAAT 563
 DB 556 ATACACATCATTTAGACACAGAGGTGAGAACGCTTACCCCGCAATCAACAGAC 615
 QY 564 TCCTTATGTGTGAATTTGTGAAGATTCAGCCGAATTTTGGGAGGTGGCGCTGTCT 623
 DB 616 ATGCTGTGTGCGACGCTGAGAGAGAGGGGCAAGAGACTCCGCAAGGTGTACTCCGGGGG 675
 QY 624 ACTGTATCTGCAAAAGCAAGCTCCAGGGAAAT-----CGAGGTGGGGCACTTC 671
 DB 676 CCTGTGTCTGTAAACAGTCTCTTCAAGGCAATATCTCTGGGGGCAAGATCCGTGTGCG 735
 QY 672 ATGGAGGGGAGGTGGCAATCTTACCAATGTTTAAATATGTAATTCCTGGATTGAAGAC 731
 DB 736 ATACCCGAAAGCCTGTGTCTTACGAAAGTCTGCAAAATATGTGATGATCCAGAG 795
 QY 732 ACTGCTAAGGACAAGT 747
 DB 796 AGCATGAAGACAATT 811

RESULT 9
 US-09-280-116-64
 Sequence 64, Application US/09280116A
 Patent No. 6331427
 GENERAL INFORMATION:
 APPLICANT: Robison, Keith E.
 TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
 FILE REFERENCE: 5800-24, 035800/176965
 CURRENT APPLICATION NUMBER: US/09/280,116A
 CURRENT FILING DATE: 1999-03-26
 NUMBER OF SEQ ID NOS: 268
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 64
 LENGTH: 452
 TYPE: DNA

ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: trypsin-like serine proteases
 US-09-280-116-64

Query Match 7.5%; Score 65; DB 4; Length 452;
 Best Local Similarity 52.0%; Pred. No. 5.4e-11;
 Matches 196; Conservative 0; Mismatches 175; Indels 6; Gaps 2;

QY 15 GAGCTAAGGGGAGAGAGCTGATCAACAGAAATATGCTTATTTGGGTGCTCCG 74
 DB 75 GAGCTAAGCTGGGCTTTCAACATCATCAATGAAGTTATCTCTGAGCCCTTGAAT 134
 QY 75 GGAACATTTTCTTTGCTGCTGATCATGCTGTTCAGAAAGAACCCCTCTCC--TATTGG 131
 DB 135 GTGACTGTGCTTTGGCTTTTATATCAGATTACAGTCAAGCTCCACTCCCTTACTTG 194
 QY 132 GTGTACTCAAGTCTCACTTCAACCCCTGTGTGGGGCTCTCATCAACCCAGCTGGTG 191
 DB 195 GTCTATTTGAATCTGACTACTTGGCCCTGCGTGGAGTCTGTATCCACCGCTTGGGTG 254
 QY 192 CTGGCCCAAGCTCACTGCTATTTTACCAATCTGAAGTATGCTGGGAAATTTCAAGAGC 251
 DB 255 ATCAACAGCTGCACACTGCAATTTACCAAGCTTCGGGTATATTGGGGGTACCAATCCA 314
 QY 252 AGAGTGAAGAGGGTACTGAACAGACA---ATTAAACCCATTCAGATGCTCGCTACTGG 308
 DB 315 GCAAGCTCTTAATGAAGAGATCTGCAGATGATTTGGTATGAGAGATGATTCATTC 374
 QY 309 AACTACAGTATAGGCCCCCAAGATGATGATCTATGCTCATCAAGCTGCTAAGCTGGCC 368
 DB 375 CACTTCTGATCACTCTATTGTATGATGATGATGATGATGATGATGATGATGATGATG 434
 QY 369 ATGCTCAATCCCAAGT 385
 DB 435 GAATCAATGACTATGT 451

RESULT 10
 US-09-386-642-10
 Sequence 10, Application US/09386642
 Patent No. 6420157
 GENERAL INFORMATION:
 APPLICANT: Dartow, Andrew
 APPLICANT: Qi, Jenson
 APPLICANT: Andrade-Gordon, Patricia
 TITLE OF INVENTION: Zymogen Activation System
 FILE REFERENCE: ORT-1028
 CURRENT APPLICATION NUMBER: US/09/386,642
 CURRENT FILING DATE: 1999-08-31
 NUMBER OF SEQ ID NOS: 60
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 10
 LENGTH: 1052
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: Fusion gene
 OTHER INFORMATION: with homo sapien serine protease catalytic domain
 US-09-386-642-10

Query Match 7.5%; Score 64.8; DB 4; Length 1052;
 Best Local Similarity 47.1%; Pred. No. 1e-10;
 Matches 290; Conservative 0; Mismatches 302; Indels 24; Gaps 2;

QY 156 CCCTGTGTGGGCGCTCTCATCAAAACCCAGCTGGGTGGTGGCCCACTCACTCTATTAA 215
 DB 241 CTCTGTGGGGCGAGCGTCAATCGCCCAAGATGGCTCTGTACAGACAGCCCACTCCCAAG 300
 QY 216 CCAATCTGAAAGTATGCTGGGAAATTTCAAGACAGAGTCAAGACGGTACTGAACAG 275
 DB 301 CCCCCTAATAGTTACCTGGGGGAGCAACCTTCAAGAGAGAGGAGGGCTGTGAGCAG 360

QY 276 ACAATTACCCATTCAGATGTCCTGCTAGCACTACAGTACAGCCCAAGAT 335
DB 361 ACCGCGACAGCACTGATGCTCTTCCACCCCGCTTCAACAACCTCCCAACAA 420
QY 336 GACC-----TCATGCTCATGAGCTGCTAAGCTGCTCACTGCTCAATCCAA 383
DB 421 GACCAACCGCATGATGATGCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGG 480
QY 384 GTCAGCCCTTCCCTGCGCAGCAGCAATGTCAGGCGAGGAGGAGGAGGAGGAG 443
DB 481 GTGAGACCCCTCACTCTCTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 444 GCTTGAATCTGAGCCCAAGAAAGAGGCGGAGACCTTGAATCTGCGGAGACCT 503
DB 541 GGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
QY 504 GCGCCGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 563
DB 601 ATCAACCATCATGAGCAACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
QY 564 TCCTTATGCTGAAATTTGTAAGTATTCAGCCGAATTTTGGGAGGAGGAGGAGG 623
DB 661 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 624 ACTGCTATCTGCAAGCAAGCTCCAGGGAAT-----CGAGTGGGAGGAGGAG 671
DB 721 CCTTGTGCTGTAACCACTCTCTCAAGGATATCTCTGAGGAGGAGGAGGAGGAG 780
QY 672 ATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 731
DB 781 ATCAACCGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 732 ACTGCTAAGGACAAGT 747
DB 841 ACGATGAAGAACATTT 856

RESULT 11
US-08-944-483-7
; Sequence 7, Application US/08944483
; Patent No. 6232456

GENERAL INFORMATION:

APPLICANT: COHEN, MAURICE
APPLICANT: COLETTIS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: KLAS, MICHAEL R.
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STEWART, KENT D.
APPLICANT: STROUPE, STEVEN D.
TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/944,483
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6183.US.01
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1166 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-944-483-7

Query Match 7.5%; Score 64.8; DB 3; Length 1166;
Best Local Similarity 47.1%; Pred.No.1.1e-10;
Matches 290; Conservative 0; Mismatches 302; Indels 24; Gaps 2;

QY 156 CCTGTGTGGGCGCTCTCATCAAAACAGCTGGGTGCTGGCCAGCTCACTGCTATTTA 215
DB 238 CTCTGTGGGCGAGCGCTCATTCGCCCAAGATGCTCTGACAGCAACCCCACTGCTCAG 227
QY 216 CCAATCTGAAATGATGCTGGGAAATTTCAAGAGCAGATCAGACGCTACTGAACAG 275
DB 298 CCGGCTACATAGTTCACTGGGGGAGCAACCTCCAGAGAGGAGGAGGAGGAGGAGGAG 357
QY 276 ACAATTACCCATTCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 335
DB 358 ACCGAGACGCACTAGTCTTCTTCCCAACCCGCTTCAACAAGCTCTCCCAACAA 417
QY 336 GACC-----TCATGCTCATCAAGCTGCTAAGCTGCTGCTGCTGCTGCTGCTGCTGCT 383
DB 418 GACCAACCGCATATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 477
QY 384 GTCAGCCCTTCTTCCCTGCGCAGCAATGTCAGGCGAGGAGGAGGAGGAGGAGGAGGAG 443
DB 478 GTGAGACCCCTCACTCTCTCTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 537
QY 444 GCTTGAATCTGAGCCCAAGAAAGAGTGGCCGACCTGCTGCTGCTGCTGCTGCTGCTGAG 503
DB 538 GGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 597
QY 504 GCGCCGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 563
DB 598 ATCAACCATCATGAGCAACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 657
QY 564 TCCTTATGCTGAAATTTGTAAGTATTCAGCCGAATTTTGGGAGGAGGAGGAGGAGGAG 623
DB 658 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717
QY 624 ACTGCTATCTGCAAGCAAGCTCCAGGGAAT-----CGAGTGGGAGGAGGAGGAGGAG 671
DB 718 CCTTGTGCTGTAACCAAGCTCTTCAAGGCAATATCTCTGAGGAGGAGGAGGAGGAGGAG 777
QY 672 ATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 731
DB 778 ATCAACCGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 837
QY 732 ACTGCTAAGGACAAGT 747
DB 838 ACGATGAAGAACATTT 853

RESULT 12
US-08-944-483-8
; Sequence 8, Application US/08944483
; Patent No. 6232456

GENERAL INFORMATION:
APPLICANT: COHEN, MAURICE
APPLICANT: COLETTIS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.

APPLICANT: GRANADOS, EDWARD N.
APPLICANT: KLAS, MICHAEL R.
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STEWART, KENT D.
APPLICANT: STROUBE, STEVEN D.
TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSER: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/944,483
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6183.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1192 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-944-483-8

Query Match
Best Local Similarity 47.1%; Score 64.8; DB 3; Length 1192;
Pred. No. 1.1e-10;
Matches 290; Conservative 0; Mismatches 302; Indels 24; Gaps 2;

156 CCTGTGTGGGCTCTCATCAAAACCAAGTGGTGTGCTGCCCCCACTCACTGCTATTTA 215
242 CTCTGTGGGCGACGCTCATCGCCCCCAAGTGGCTCTGACAGACCCCACTGCTCAAG 301
216 CCAATCTGAAGTGTGCTGGAATTTCAAGAGCAGAGTGTGAGACGCTACTGAACAG 275
302 CCCCCCTACATGATGACCTGCGGGGACAGACCACTCCAGAGAGAGGGCTGTGACAG 361
276 ACAATTAAACCCATTGATGATGTCGCTACTGAGACTACAGTATGACGCCCCACAGAGT 335
362 ACCCGGACAGCCACTAGTCTCTTCCCCCAACCCGCTTCAACAACAGCTCCCCCAACAA 421
336 GACC-----TCATGCTCATCAAGCTGGCTAAGCTGCCCTCCATGCTCATATCCCAA 383
422 GACCCAGCGAATGACATCATGCTGTGAGATGGCATGCGCACTTCCATCACTGGCGCT 481
384 GTCCAGCCCTTCTCCCTGCGCACCAACATGTGAGGCGAGGCACTGTCTGTACTCTCA 443
482 GTGCGAGCCCTTCACTCTCTCAGCGTGTGTACGTGTGACGACAGCTGCCCTCATTTTC 541
444 GGTGTGACTGAGCCAGAAAGAAACAGTGGCGACACCTGACTTGGCGGCAAACTGGAG 503
542 GGCTGGGGGAGCAGTTCACAGCCCGGAGTTACGCTGCTCAACACTTGGGATGCGCAAC 601

504 GCCCCGATGATCTGATGAGAAATGCAAAACAGAAACAGAAACCAAGCAAGAT 563
602 ATCACCATTATGAGACACAGAAAGTGTGAAAGCGCTTACCCGCAACATCAAGACAC 661
564 TCCTATGTGTGAATTTGTGAAGTATTACAGCCCAATTTTGGGAGGTGGCGCTTCT 623
662 ATGCTGTGCGCAGCGCTGAGAGAGGGGCAAGAGACTCTGCGCAGGATGATCTCCGGGG 721
624 ACTGATCTGCAAGACAGAGCTCCAGGAT-----CGAGTGGGGCACTTC 671
722 CCTGTGTGTAAACAGTCTCTTCAAGCAATATCTCTGGGGCAAGATCCGTGTGCG 781
672 ATGGAGGGAGAGTGGCATCTTACACCAATGTTTAAATATGATCTGATTTGAGAAC 731
782 ATCACCAGAAAGCTGTGTCTTACAGAAAGTGTGAATATGTGATGATCCAGAG 841
732 ACTGTAAGACAGT 747
842 ACGATGAGAGCAATT 857

RESULT 13
US-09-025-059-2
Sequence 2, Application US/09025059
Patent No. 6075136
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: Guegler, Karl J.
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/025,059
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0481 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1314 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LUNGUT10
CLONE: 2723646
US-09-025-059-2

Query Match
Best Local Similarity 47.1%; Score 64.8; DB 3; Length 1314;
Pred. No. 1.2e-10;
Matches 290; Conservative 0; Mismatches 302; Indels 24; Gaps 2;

```

QY 156 CCCTGTGTGGCGCTCTCATCAAAACCAGTGGGTGCTGGCCCCAGCTCACTGCTATTTA 215
D 359 CTCTGTGTGGCGGAGCGCTCATTCGCCCCAGATGGTCTCTGACAGAGCCCACTGGCTCAAG 418
QY 216 CCAATCTGAAATGATGCTGGGAAATTTTCAAGAGCAGTCAAGACGGTACTGAAACAG 275
D 419 CCCCCGTACATAGTTCACCTGGGGCAGCAACCTCCAGAAAGAGGGGCTGTGAGCAG 478
QY 276 ACAATTAACCCCATTCAGATGCTCCGCTACTGAAATACATGATCAAGCCCCCAGAGAT 335
D 479 ACCCGAGACAGCAGTGTGCTCTTCCCGACCCGGCTTCAACAACAGCTTCCCAACAA 538
QY 336 GACC-----TCATGCTCATCAAGCTGGCTGAAGCTGGCTGCTCAATCCCAA 383
D 539 GACCACCGCAATGACATCATGTGTGTAAGATGGCATGCGCAATCTCATCACTGGGCT 598
QY 384 GTCCAGGCTCTTCCCTGCGCACCAATGTCAAGGCGAGCACTGTCTGTACTCTCA 443
D 599 GTCCGACCCCTCAACCTCTCTCTCAAGCTGTGTCACTGTGGCACACAGCTGCTCATTTCC 658
QY 444 GGTTTGACTGAGCCCAAGAAACAGTGGCCGACACCTGACTTGGGGCAGAACTGGAG 503
D 659 GCGTGGGGCAGCAGCTCAACCCCACTTACGCTGCTCAACCTTGCATGCGCCAC 718
QY 504 GCCCCGCTGATGCTGATGCAATGCAAAACAGAAAGAAAGAAAGCAAGAAAT 563
D 719 ATCACCATCATTTAGACACCAAGTGTGAGAGAGCTTACCCCGCAACATCAAGACAC 778
QY 564 TCCTTATGTGTGAATTTGTGAAAGTATTCAGCCGAATTTTGGGGAGTGGCCGTGCT 623
D 779 ATGATGTGTCCAGCGTGCAGAAAGGGGGCAGAGACTCTGCCAGGCTGACTCCGGGGGC 838
QY 624 ACTGTCATGCAAGACAGAGCTCCAGGGAAT-----CGAGTGGGGCACTTC 671
D 839 CCTGTGCTGTGAACAGTCTCTTCAAGGCAATATCTCTGGGGCCAGAACTCGTGGC 898
QY 672 ATGGAGGGGAGCTCGGATCTACCAATGTTTACAATATGATCTGATTTAGAAC 731
D 899 ATACCCGAAAGGCTGGTGTCTACAGAAAGTCTGAAATATGTGACTGTGATCCAGAG 958
QY 732 ACTGCTAAGACAAAT 747
D 959 ACGATGAAGAACAAAT 974

```

RESULT 14
 US-08-790-137-2
 ; Sequence 2, Application US/08790137
 ; Patent No. 5840871
 ; GENERAL INFORMATION:
 ; APPLICANT: Hillman, Jennifer L.
 ; TITLE OF INVENTION: A NOVEL PROSTATE-ASSOCIATED
 ; TITLE OF INVENTION: KALIKREIN
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/790,137
 ; FILING DATE: Filed Herewith
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:

```

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0195 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-845-4166
; TELEFAX: 415-845-0555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 833 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-790-137-2

Query Match 7.3%; Score 63.4; DB 2; Length 833;
Best Local Similarity 46.8%; Pred. No. 2, 6e-10;
Matches 288; Conservative 0; Mismatches 304; Indels 24; Gaps 2;

QY 156 CCCTGTGTGGCGCTCTCATCAAAACCAGTGGGTGCTGGCCCCAGCTCACTGCTATTTA 215
D 171 CTCTGTGTGGCGGAGCGCTCATTCGCCCCAGATGGTCTCTGACAGAGCCCACTGGCTCAAG 220
QY 216 CCAATCTGAAATGATGCTGGGAAATTTTCAAGAGCAGTCAAGACGGTACTGAAACAG 275
D 231 CCCCCGTACATAGTTCACCTGGGGCAGCAACCTCCAGAAAGAGGGGCTGTGAGCAG 290
QY 276 ACAATTAACCCCATTCAGATGCTCCGCTACTGGAATCAAGTATAGCCGCCACAGAT 335
D 291 ACCCGAGACAGCAGTGTGCTTCCCGACCCCGGCTTCAACAAGAGCTTCCCAACAA 350
QY 336 GACC-----TCATGCTCATCAAGCTGGCTGAAGCTGGCTGCTCAATCCCAA 383
D 351 GACCACCGCAATGACATCATGTGTGTAAGATGGCATGCGCAATCTCATCACTGGGCT 410
QY 384 GTCCAGGCTCTTCCCTGCGCACCAATGTCAAGGCGAGCACTGTCTGTACTCTCA 443
D 411 GTCCGACCCCTCAACCTCTCTCTCAAGCTGTGTCACTGTGGCACACAGCTGCTCATTTCC 470
QY 444 GGTTTGACTGAGCCCAAGAAACAGTGGCCGACACCTGACTTGGGGCAGAACTGGAG 503
D 471 GCGTGGGGCAGCAGCTCAACCCCACTTACGCTGCTCAACCTTGGCATGCGCCAC 530
QY 504 GCCCCGCTGATGCTGATGCAATGCAAAACAGAAAGAAAGAAAGCAAGAAAT 563
D 531 ATCACCATCATTTAGACACCAAGTGTGAGAGAGCTTACCCCGCAACATCAAGACAC 590
QY 564 TCCTTATGTGTGAATTTGTGAAAGTATTCAGCCGAATTTTGGGGAGTGGCCGTGCT 623
D 591 ATGATGTGTCCAGCGTGCAGAAAGGGGGCAGAGACTCTGCCAGGGTACTCGGGGGC 650
QY 624 ACTGTCATGCAAGACAGTCTCCAGGGAAT-----CGAGTGGGGCACTTC 671
D 651 CCTGTGCTGTGAACAGTCTCTTCAAGGCAATATCTCTGGGGCCAGAACTCGTGGC 710
QY 672 ATGGAGGGGAGCTCGGATCTACCAATGTTTACAATATGATCTGATTTAGAAC 731
D 711 ATCACCAGAAAGCTGTGTCTACAGAAAGTGTGAAATATGTGACTGTGATCCAGAG 770
QY 732 ACTGCTAAGACAAAT 747
D 771 ACGATGAAGAACAAAT 786

```

RESULT 15
 US-08-361-395-2
 ; Sequence 2, Application US/08361395
 ; Patent No. 5733768
 ; GENERAL INFORMATION:
 ; APPLICANT: Dixon, Eric P.
 ; APPLICANT: Johnstone, Edward M.
 ; APPLICANT: Little, Sheila P.

APPLICANT: No. 57337681s, Franklin H.
 TITLE OF INVENTION: PROTEASE AND RELATED DNA COMPOUNDS
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Eli Lilly and Company
 STREET: Lilly Corporate Center/Patent Division
 CITY: Indianapolis
 STATE: Indiana
 COUNTRY: United States
 ZIP: 46285
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/361,395
 FILING DATE: 22-DEC-1994
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Blalock, Donna K.
 REGISTRATION NUMBER: 38,082
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 317-277-1090
 TELEFAX: 317-276-3861
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 732 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-08-361-395-2

Query Match 7.3%; Score 63; DB 1; Length 732;
 Best Local Similarity 53.4%; Pred. No. 3.2e-10;
 Matches 132; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY	156	CCCTGTGTGGGCTCTCTCATCAACCAGCTGGGTGCTGCCCCAGCTCAGTCTATTTA	215
DB	136	CTCTGTGTGGGGGTCTTATCCATCCATGTGGGTCTCTCACAGCTGCCCAAAAAA	195
QY	216	CCAAATCTGAAAGTGAATGCTGGGAAATTTCAGAGCAGAGTCAGAGCGCTACTGAACAG	275
DB	196	CCGAATCTTCAGGTCTTCTCTGGGGAAGCATACCTTCGGCAAGGAGAGTTCACAGAG	255
QY	276	ACAATTAAACCCATTGATCGTCCGCTACTGGAATCAAGTCAATAGCGCCACAGAGAT	335
DB	256	CAGAGTTCGTGTGTCGGGGGTGTGATCCACCTGACTATGATGCCGACCATGACAG	315
QY	336	GACCTCATGCTCATCAAGCTGCTAAAGCTGCATGCTCAATCCAAAGTCCAGCCCTT	395
DB	316	GACATCATGCTGTGGCGCTGGCAGCCGACCAAACTCTGAACATCATCAGCCCTT	375
QY	396	CCCTCG 402	
DB	376	CCCTGG 382	

Search completed: November 7, 2003, 23:05:40
 Job time: 84 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2003, 23:04:21 ; Search time 351 Seconds
(without alignments)
7872.847 Million cell updates/sec

Title: US-10-037-270-482

Perfect score: 866
Sequence: 1 ggcactactccctgagcta.....tgggaatgacatagctagt 866

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 214354 seqs, 1595478879 residues

Total number of hits satisfying chosen parameters: 4282708

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	866	100.0	866	14	US-10-037-270-482
2	862.8	99.6	1060	10	US-09-910-151-1
3	850.8	98.2	865	12	US-10-246-006-1
4	730	84.3	730	11	US-09-842-758-11
5	728.4	84.1	730	11	US-09-842-758-13
6	706.4	81.6	708	10	US-09-810-151-3
7	703.8	81.3	721	11	US-09-842-758-106
8	606.4	70.0	798	9	US-09-888-615-43
9	511.2	59.0	705	12	US-10-246-006-3
10	154.4	17.8	853	10	US-10-101-510-199
11	154	17.8	850	12	US-09-954-455-748
12	153.2	17.7	802	9	US-09-923-779-148
13	153.2	17.7	802	12	US-10-101-510-210
14	142	16.4	699	12	US-10-181-808-1
15	132.8	15.3	522	9	US-09-923-779-91
16	132.8	15.3	534	9	US-09-923-779-97

C 17	132.8	15.3	566	9	US-09-923-779-98	Sequence 98, Appl
C 18	132.8	15.3	567	9	US-09-923-779-142	Sequence 142, App
C 19	132.8	15.3	572	9	US-09-923-779-21	Sequence 21, Appl
C 20	132.8	15.3	592	13	US-10-040-739-1120	Sequence 1120, Ap
C 21	131.8	15.2	566	9	US-09-923-779-99	Sequence 99, Appl
C 22	131	15.1	556	9	US-09-923-779-86	Sequence 86, Appl
C 23	130.6	15.1	565	9	US-09-923-779-92	Sequence 92, Appl
C 24	130.4	15.1	566	9	US-09-923-779-100	Sequence 100, Appl
C 25	130.2	15.0	560	9	US-09-923-779-83	Sequence 83, Appl
C 26	120.8	13.9	567	14	US-10-060-036-4206	Sequence 4206, Ap
C 27	102	11.8	957	12	US-10-170-789-51	Sequence 51, Appl
C 28	102	11.8	957	13	US-10-045-367A-1	Sequence 1, Appl1
C 29	102	11.8	1123	12	US-10-137-870-247	Sequence 247, App
C 30	102	11.8	1123	12	US-10-140-018-247	Sequence 247, App
C 31	102	11.8	1123	12	US-10-140-021-247	Sequence 247, App
C 32	102	11.8	1123	12	US-10-140-274-247	Sequence 247, App
C 33	102	11.8	1123	12	US-10-140-471-247	Sequence 247, App
C 34	102	11.8	1123	12	US-10-140-807-247	Sequence 247, App
C 35	102	11.8	1123	12	US-10-140-922-247	Sequence 247, App
C 36	102	11.8	1123	12	US-10-140-924-247	Sequence 247, App
C 37	102	11.8	1123	12	US-10-140-925-247	Sequence 247, App
C 38	102	11.8	1123	12	US-10-141-698-247	Sequence 247, App
C 39	102	11.8	1123	12	US-10-141-702-247	Sequence 247, App
C 40	102	11.8	1123	12	US-10-141-704-247	Sequence 247, App
C 41	102	11.8	1123	12	US-10-142-421-247	Sequence 247, App
C 42	102	11.8	1123	12	US-10-142-433-247	Sequence 247, App
C 43	102	11.8	1123	12	US-10-142-767-247	Sequence 247, App
C 44	102	11.8	1123	12	US-10-143-033-247	Sequence 247, App
C 45	102	11.8	1123	12	US-10-144-994-247	Sequence 247, App

ALIGNMENTS

RESULT 1
US-10-037-270-482
Sequence 482, Application US/10037270
GENERAL INFORMATION: US20030104529A1
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyun
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aifong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungting
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: Tillinhuast, John
APPLICANT: Drmanac, Radost T.
TITLE OF INVENTION: No. US20030104529A1 Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP28
CURRENT APPLICATION NUMBER: US/10/037,270
CURRENT FILING DATE: 2002-01-04
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: pc_fl_genes Version 1.0
SEQ ID NO 482
LENGTH: 866
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS

LOCATION: (42) .. (749)
US-10-037-270-482

Query Match 100.0%; Score 866; DB 14; Length 866;
Best Local Similarity 100.0%; Pred. No. 2,76-278;
Matches 866; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GGCACCTACTCCCGAGCTAAGGGGGAAGAGCTGATCACCAGTAATATGCTTCTATT 60
DB 1 GGCACCTACTCCCGAGCTAAGGGGGAAGAGCTGATCACCAGTAATATGCTTCTATT 60
QY 61 TGGGTGCTCTGCTGGAGCATTTTCTTGTGATCTGATCTGTTCAAGAAAGACCTTG 120
DB 61 TGGGTGCTCTGCTGGAGCATTTTCTTGTGATCTGATCTGTTCAAGAAAGACCTTG 120
QY 121 CTCCCTATTTGGTGTACTTCAAGCTCACTTCAACCCCTGTGTGGGGTCTCTATCAAC 180
DB 121 CTCCCTATTTGGTGTACTTCAAGCTCACTTCAACCCCTGTGTGGGGTCTCTATCAAC 180
QY 181 CCAGCTGGGTGTGGGCCAGCTGCTGCTATTACCAAAATCTGAAGTGTGCTGGGAA 240
DB 181 CCAGCTGGGTGTGGGCCAGCTGCTGCTATTACCAAAATCTGAAGTGTGCTGGGAA 240
QY 241 ATTTCAAGAGAGAGTGAAGAGCGGTACTGAACAGACATTTAACCCATTGAGTCGTCC 300
DB 241 ATTTCAAGAGAGAGTGAAGAGCGGTACTGAACAGACATTTAACCCATTGAGTCGTCC 300
QY 301 GCTACTGTGAAGTACAGTCTATGCCCCCAGAGATGACCTGATCTCATAGCTGGCTA 360
DB 301 GCTACTGTGAAGTACAGTCTATGCCCCCAGAGATGACCTGATCTCATAGCTGGCTA 360
QY 361 AGCTGCGATGCTCAATCCCAAGTCCAGCCCTTCCCTGCGACCAACCAATGTCAGGC 420
DB 361 AGCTGCGATGCTCAATCCCAAGTCCAGCCCTTCCCTGCGACCAACCAATGTCAGGC 420
QY 421 CAGGACCTGTCTGTCTACTCTGAGTTTGAAGCTGAGCCAAAGAAAGAGTGGCCGAC 480
DB 421 CAGGACCTGTCTGTCTACTCTGAGTTTGAAGCTGAGCCAAAGAAAGAGTGGCCGAC 480
QY 481 CTGACTTGGCGGAGAACTGAGAGCCCCCGTGTGATCTGATCGAATGCCAAAAAAG 540
DB 481 CTGACTTGGCGGAGAACTGAGAGCCCCCGTGTGATCTGATCGAATGCCAAAAAAG 540
QY 541 AACAAAGAAAAAGCCACAGAAATTCCTTATGTGAAATTTGTAAAGTATTCAGCCGAA 600
DB 541 AACAAAGAAAAAGCCACAGAAATTCCTTATGTGAAATTTGTAAAGTATTCAGCCGAA 600
QY 601 TTTTGGGAGAGTGGCGGTGCTACTGTCTGCAAGCAAGCTCCAGGGAATCGAGG 660
DB 601 TTTTGGGAGAGTGGCGGTGCTACTGTCTGCAAGCAAGCTCCAGGGAATCGAGG 660
QY 661 TGGGGCACTTCAATGAGGAGGAGCGTCGCAATCTACCAATGTTTAAATATGATCT 720
DB 661 TGGGGCACTTCAATGAGGAGGAGCGTCGCAATCTACCAATGTTTAAATATGATCT 720
QY 721 GGATTGGAACACTGCTAAGGAAAGTGAAGCCCTACTCTCCCTGCAATTCACATGGC 780
DB 721 GGATTGGAACACTGCTAAGGAAAGTGAAGCCCTACTCTCCCTGCAATTCACATGGC 780
QY 781 TCTGCACTGAGCTATACAGAGATAATTTTCCCTATTCAAATAATATCTCAATG 840
DB 781 TCTGCACTGAGCTATACAGAGATAATTTTCCCTATTCAAATAATATCTCAATG 840
QY 841 AAAATTTGGGAATGTACATCTAGT 866
DB 841 AAAATTTGGGAATGTACATCTAGT 866
```

RESULT 2
US-09-910-151-1
Sequence 1, Application US/09910151
Patent No. US20020137181A1
GENERAL INFORMATION:

```
APPLICANT: Meyers, Rachel
TITLE OF INVENTION: 14087, A NOVEL SERINE PROTEASE MOLECULE
FILE REFERENCE: 38155-20021.00
CURRENT APPLICATION NUMBER: US/09/910.151
PRIOR FILING DATE: 2001-07-18
PRIORITY DATE: 2000-07-17
NUMBER OF SEQ. ID NOS: 11
SOFTWARE: PASTSEQ for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1060
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (164)...(871)
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(1060)
OTHER INFORMATION: n = A,T,C or G
US-09-910-151-1
```

Query Match 99.6%; Score 862.8; DB 10; Length 1060;
Best Local Similarity 99.8%; Pred. No. 3,6e-277;
Matches 864; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 GGCACCTACTCCCGAGCTAAGGGGGAAGAGCTGATCACCAGTAATATGCTTCTATT 60
DB 123 GACACTTACTCCCGAGCTAAGGGGGAAGAGCTGATCACCAGTAATATGCTTCTATT 182
QY 61 TGGGTGCTCTGCTGGAGCATTTTCTTGTGATCTGATCTGTTCAAGAAAGACCTTG 120
DB 183 TGGGTGCTCTGCTGGAGCATTTTCTTGTGATCTGATCTGTTCAAGAAAGACCTTG 242
QY 121 CTCCCTATTTGGTGTACTTCAAGCTCACTTCAACCCCTGTGTGGGGTCTCTATCAAC 180
DB 243 CTCCCTATTTGGTGTACTTCAAGCTCACTTCAACCCCTGTGTGGGGTCTCTATCAAC 302
QY 181 CCAGCTGGGTGTGGGCCAGCTGCTGCTATTACCAAAATCTGAAGTGTGCTGGGAA 240
DB 303 CCAGCTGGGTGTGGGCCAGCTGCTGCTATTACCAAAATCTGAAGTGTGCTGGGAA 362
QY 241 ATTTCAAGAGAGAGTGAAGAGCGGTACTGAACAGACATTTAACCCATTGAGTCGTCC 300
DB 363 ATTTCAAGAGAGAGTGAAGAGCGGTACTGAACAGACATTTAACCCATTGAGTCGTCC 422
QY 301 GCTACTGTGAAGTACAGTCAATAGGCCCCACAGAGTGAACCTGATCTATCAAGCTGGCTA 360
DB 423 GCTACTGTGAAGTACAGTCAATAGGCCCCACAGAGTGAACCTGATCTATCAAGCTGGCTA 482
QY 421 CAGGACCTGTCTGTCTACTCTGAGTTTGAAGCTGAGCCAAAGAAAGAGTGGCCGAC 480
DB 543 CAGGACCTGTCTGTCTACTCTGAGTTTGAAGCTGAGCCAAAGAAAGAGTGGCCGAC 602
QY 541 AACAAAGAAAAAGCCACAGAAATTCCTTATGTGAAATTTGTAAAGTATTCAGCCGAA 600
DB 663 AACAAAGAAAAAGCCACAGAAATTCCTTATGTGAAATTTGTAAAGTATTCAGCCGAA 722
QY 601 TTTTGGGAGAGTGGCGGTGCTACTGCTATCGCAAGCAAGCTCCAGGGAATCGAGG 660
DB 723 TTTTGGGAGAGTGGCGGTGCTACTGCTATCGCAAGCAAGCTCCAGGGAATCGAGG 782
QY 661 TGGGGCACTTCAATGAGGAGGAGCGTCGCAATCTACCAATGTTTAAATATGATCTCT 720
```

Db 783 TGGGGCACTTCATGAGGGGAGCGTCGCATCTACACCAATGTTTACAAATATGATCTCT 842
Qy 721 GGATTGAGAACACTGCTAAGGACAAAGTGAAGCCCTACTTCTCCCTGCAATTCACCTGGC 780
Db 843 GGATTGAGAACACTGCTAAGGACAAAGTGAAGCCCTACTTCTCCCTGCAATTCACCTGGC 902
Qy 781 TCTGCCATGACCTATACAGAGCAATATTTTCCCTCTATTCAAATTAATCTCCAAATG 840
Db 903 TCTGCCATGACCTATACAGAGCAATATTTTCCCTCTATTCAAATTAATCTCCAAATG 962
Qy 841 AAAATTTGGGAATGTAGCATCTACT 866
Db 963 AAAATTTGGGAATGTAGCATCTACT 988

RESULT 3
US-10-246-006-1
; Sequence 1, Application US/10246006
; Publication No. US20030144493A1
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; TITLE OF INVENTION: Human Serine Protease
; FILE REFERENCE: 00-16
; CURRENT APPLICATION NUMBER: US/10/246,006
; CURRENT FILING DATE: 2002-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 865
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (42)...(746)
US-10-246-006-1

Query Match 98.2%; Score 850.8; DB 12; Length 865;
Best Local Similarity 99.7%; Pred. No. 3,3e-273;
Matches 863; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 1 GGCACTTACTCTCCCTGAGCTAAGGGGGAAGAGCTGATCACCAGTAAATATGCTCTTCTATT 60
Db 1 GGCACTTACTCTCCCTGAGCTAAGGGGGAAGAGCTGATCACCAGTAAATATGCTCTTCTATT 60
Qy 61 TGGGTGCTCTGCTGGGACATTTTCTTTGCTGACTCATCTGTTCAAGAAAGAACCCCTG 120
Db 61 TGGGTGCTCTGCTGGGACATTTTCTTTGCTGACTCATCTGTTCAAGAAAGAACCCCTG 120
Qy 121 CTCCTTATTTGGTGTACTCAAGTCTCACTCAACCCCTGTGTGGGGTCTCTATCAAC 180
Db 121 CTCCTTATTTGGTGTACTCAAGTCTCACTCAACCCCTGTGTGGGGTCTCTATCAAC 180
Qy 181 CCAGCTGGGTGCTGGCCCGCACTCATCTATTTACCAATCTGAAAGTGAATGCTGGGAA 240
Db 181 CCAGCTGGGTGCTGGCCCGCACTCATCTATTTACCAATCTGAAAGTGAATGCTGGGAA 240
Qy 241 ATTTCAGAGAGAGTCAAGAGCGTACTGAACAGCAATTAACCCATTCAATCTGCTCC 300
Db 241 ATTTCAGAGAGAGTCAAGAGCGTACTGAACAGCAATTAACCCATTCAATCTGCTCC 300
Qy 301 GCTACTGGAAGTCAAGTCAAGCGCCCAAGATGACCTCATGCTCATCAAGCTGGCTA 360
Db 301 GCTACTGGAAGTCAAGTCAAGCGCCCAAGATGACCTCATGCTCATCAAGCTGGCTA 360
Qy 361 AGCTGTCATGCTCAATCCCAAGTCCAGCCCTTCCCTGCGACCAACCAATGCTAGGC 420
Db 361 AGCTGTCATGCTCAATCCCAAGTCCAGCCCTTCCCTGCGACCAACCAATGCTAGGC 420
Qy 421 CAGGACCTGTCTGTCTACTCTCAAGTTTGAAGTGAAGCAAGAAAGAGTGGCGACACC 480
Db 421 CAGGACCTGTCTGTCTACTCTCAAGTTTGAAGTGAAGCAAGAAAGAGTGGCGACACC 480
Qy 481 CTGACTTGCGGCAAGACCTGAGGCGCCCGTGTATGTCTGATGAGAAATGCCAAAAACAG 540

Db 481 CTGACTTGCGGCAAGACCTGAGGCGCCCGTGTATGTCTGATGAGAAATGCCAAAAACAG 540
Qy 541 AACAGAGAAAACCCACAGAAATTTCTTATGTTGAAATTTGGAAGTATTCAGCCGAA 600
Db 541 AACAGAGAAAACCCACAGAAATTTCTTATGTTGAAATTTGGAAGTATTCAGCCGAA 600
Qy 601 TTTTGGGAGAGTGGCCGTGCTACTGTCATCTGCAAGACCAAGCTCCAGGAATCGAG 660
Db 601 TTTTGGGAGAGTGGCCGTGCTACTGTCATCTGCAAGACCAAGCTCCAGGAATCGAG 660
Qy 661 TGGGGCACTTCATGAGAGGGGAGCTGGGACATCTACACCAATGTTTCAAAATATGATCT 720
Db 661 TGGGGCACTTCATGAGAGGGGAGCTGGGACATCTACACCAATGTTTCAAAATATGATCT 720
Qy 721 GGATTGAGAACACTGCTAAGGACAAAGTGAAGCCCTACTTCTCCCTGCAATTCACCTGGC 780
Db 721 GGATTGAGAACACTGCTAAGGACAAAGTGAAGCCCTACTTCTCCCTGCAATTCACCTGGC 780
Qy 781 TCTGCCATGACCTATACAGAGCAATATTTTCCCTCTATTCAAATTAATCTCCAAATG 840
Db 781 TCTGCCATGACCTATACAGAGCAATATTTTCCCTCTATTCAAATTAATCTCCAAATG 840
Qy 841 AAAATTTGGGAATGTAGCATCTACT 866
Db 841 AAAATTTGGGAATGTAGCATCTACT 866

RESULT 4
US-09-842-758-11
; Sequence 11, Application US/09842758
; Publication No. US20030083244A1
; GENERAL INFORMATION:
; APPLICANT: Vermet, Corine A. M.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Shinkels, Richard A.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Zernusen, Bryan D.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Majumder, Kumud
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Paturajan, Meeta
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gangoli, Saha A.
; APPLICANT: Smithson, Glenda
; APPLICANT: Rastelli, Luca
; APPLICANT: Macdougall, John R.
; APPLICANT: Teapler, Raymond J.
; APPLICANT: Grose, William M.
; APPLICANT: Edward, Szekeres S.
; APPLICANT: Alsobrook II, John P.
; TITLE OF INVENTION: No. US20030083244A1e1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-783
; CURRENT APPLICATION NUMBER: US/09/842,758
; CURRENT FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/200,158
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/200,613
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,780
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/201,006
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,007
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,236
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,238
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,186

PRIOR FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: 60/201,474
PRIOR FILING DATE: 2000-05-03
PRIOR APPLICATION NUMBER: 60/201,508
PRIOR FILING DATE: 2000-05-03
PRIOR APPLICATION NUMBER: 60/220,591
PRIOR FILING DATE: 2000-07-25
PRIOR APPLICATION NUMBER: 60/232,678
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 60/263,217
PRIOR FILING DATE: 2001-01-22
PRIOR APPLICATION NUMBER: 60/265,160
PRIOR FILING DATE: 2001-01-30
NUMBER OF SEQ ID NOS: 113
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 730
TYPE: DNA
ORGANISM: Homo sapiens
US-09-842-758-11

Query Match 84.3%; Score 730; DB 11; Length 730;
Best Local Similarity 100.0%; Pred. No. 7e-233;

Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 35 GATCACCATGAATATGCTCTATTTGGTGTCTGCTGGGACATTTTCTTTGCTGA 94
DB 1 GATCACCATGAATATGCTCTATTTGGTGTCTGCTGGGACATTTTCTTTGCTGA 60
OY 95 CTCTATCTGTGAGAAAGAACCCCTGCTCTCTATTTGGTGTGATCTCAAGTCTCACTTCAA 154
DB 61 CTCTATCTGTGAGAAAGAACCCCTGCTCTCTATTTGGTGTGATCTCAAGTCTCACTTCAA 120
OY 155 CCCCCTGTGGGCGCTCTCATCAAAACCAAGTGGGTGTGGCCCCCACTGCTGCTATT 214
DB 121 CCCCCTGTGGGCGCTCTCATCAAAACCAAGTGGGTGTGGCCCCCACTGCTGCTATT 180
OY 215 ACCAAATCTGAAGTGTGCTGGGAAATTTCAAGAGCAGATCTAGAGAGCGGTACTGAACA 274
DB 181 ACCAAATCTGAAGTGTGCTGGGAAATTTCAAGAGCAGATCTAGAGAGCGGTACTGAACA 240
OY 275 GACAATTAACCCCATTCAGATGTCCTGCTACTGGAATCAAGTCAATAGCCGCCACAGAGA 334
DB 241 GACAATTAACCCCATTCAGATGTCCTGCTACTGGAATCAAGTCAATAGCCGCCACAGAGA 300
OY 335 TGACCTCATGCTCATCAAGCTGGGCTAAGCTGCTCAATGCCAAAGTCCAGCCCT 394
DB 301 TGACCTCATGCTCATCAAGCTGGGCTAAGCTGCTCAATGCCAAAGTCCAGCCCT 360
OY 395 TCCCTCGCCACCAATGTCAGGCGCAGGCACTGCTGCTACTCTCAGGTTTGAATG 454
DB 361 TCCCTCGCCACCAATGTCAGGCGCAGGCACTGCTGCTACTCTCAGGTTTGAATG 420
OY 455 GAGCCAAAGAAAAGTGGCCGACACCTGACTTGGCGAGAACTGAGAGCCCGCTGAT 514
DB 421 GAGCCAAAGAAAAGTGGCCGACACCTGACTTGGCGAGAACTGAGAGCCCGCTGAT 480
OY 515 GTCTGATGAGAAATGCCCCAAAACAGAAAGAAAAAGCCACAGAAATCTTATATGCT 574
DB 481 GTCTGATGAGAAATGCCCCAAAACAGAAAGAAAAAGCCACAGAAATCTTATATGCT 540
OY 575 GAAATTTGTGAAGATATCAGCCGAATTTTGGGAGGTGGCCGTTGCTACTGTCATCTG 634
DB 541 GAAATTTGTGAAGATATCAGCCGAATTTTGGGAGGTGGCCGTTGCTACTGTCATCTG 600
OY 635 CAAAGACAAGCTCCAGGGAATCGAGTGGGCACTTCATGAGAGGGAAGTCCGCACTTA 694
DB 601 CAAAGACAAGCTCCAGGGAATCGAGTGGGCACTTCATGAGAGGGAAGTCCGCACTTA 660
OY 695 CACCAATGTTTAAATATGATCTTGATGAGAAACCTGCTAAGAGCAAGTGAAGACC 754
DB 661 CACCAATGTTTAAATATGATCTTGATGAGAAACCTGCTAAGAGCAAGTGAAGACC 720

OY 755 TACTTCTCCC 764
DB 721 TACTTCTCCC 730

RESULT 5

US-09-842-758-13
Sequence 13, Application US/09842758
Publication No. US20030083244A1

GENERAL INFORMATION:
APPLICANT: Vernet, Corine A. M.
APPLICANT: Fernandez, Elma R.
APPLICANT: Gerlach, Valerie
APPLICANT: Shinkels, Richard A.
APPLICANT: Malyankar, Uriel M.
APPLICANT: Boldog, Ferenc L.
APPLICANT: Zernusen, Bryan D.
APPLICANT: Spytek, Kimberly A.
APPLICANT: Majumder, Kumud
APPLICANT: Tchernev, Velizar T.
APPLICANT: Padigaru, Muralidhara
APPLICANT: Baturajan, Meera
APPLICANT: Burgess, Catherine E.
APPLICANT: Gangoli, Esha A.
APPLICANT: Smithson, Glenda
APPLICANT: Raestelli, Luca
APPLICANT: MacDougall, John R.
APPLICANT: Taupier, Raymond J.
APPLICANT: Grose, William M.
APPLICANT: Edward, Szekeres S.
APPLICANT: Alsobrook II, John P.
TITLE OF INVENTION: No. US20030083244A1 Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 15966-783
CURRENT APPLICATION NUMBER: US/09/842,758
CURRENT FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: 60/200,158
PRIOR FILING DATE: 2000-04-26
PRIOR APPLICATION NUMBER: 60/200,613
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,780
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/201,006
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,007
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,236
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,238
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,186
PRIOR FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: 60/201,474
PRIOR FILING DATE: 2000-05-03
PRIOR APPLICATION NUMBER: 60/201,508
PRIOR FILING DATE: 2000-05-03
PRIOR APPLICATION NUMBER: 60/220,591
PRIOR FILING DATE: 2000-07-25
PRIOR APPLICATION NUMBER: 60/232,678
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 60/263,217
PRIOR FILING DATE: 2001-01-22
PRIOR APPLICATION NUMBER: 60/265,160
PRIOR FILING DATE: 2001-01-30
NUMBER OF SEQ ID NOS: 113
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 730
TYPE: DNA
ORGANISM: Homo sapiens
US-09-842-758-13

Query Match 84.1%; Score 728.4; DB 11; Length 730;
Best Local Similarity 99.9%; Pred. No. 2.4e-232;

Matches 729; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 35 GATCACCATGAAATATGCTCTATTTGGGTGCTCGCTGGACATTTTCTTGCTGA 94
DB 1 GATACCATGAAATATGCTCTATTTGGGTGCTCGCTGGACATTTTCTTGCTGA 60
QY 95 CTGATCTGTTGAGAAAGAACCTGCTCCCTATTTGGTGTACCTCAAGTCTCACTTCAA 154
DB 61 CTGATCTGTTGAGAAAGAACCTGCTCCCTATTTGGTGTACCTCAAGTCTCACTTCAA 120
QY 155 CCCCCTGTGGGCGCTCTCAATCAAAACCAAGTGGGTGTGGCCCCAGCTCACTCTATT 214
DB 121 CCCCCTGTGGGCGCTCTCAATCAAAACCAAGTGGGTGTGGCCCCAGCTCACTCTATT 180
QY 215 ACCAAATCTGAAATGATGCTGGGAAATTTCAAGAGCAGATCAGAGCGGTATGAAACA 274
DB 181 ACCAAATCTGAAATGATGCTGGGAAATTTCAAGAGCAGATCAGAGCGGTATGAAACA 240
QY 275 GACAATTAACCCCATTCAGATCGTCCGCTACTGGAATCAAGTCAATAGCCGCCACAGAGA 334
DB 241 GACAATTAACCCCATTCAGATCGTCCGCTACTGGAATCAAGTCAATAGCCGCCACAGAGA 300
QY 335 TGACCTCATGCTCATCAAGCTGGCTAAGCTGCGCATGCTCAATCCCAAGTCCAGCCCT 394
DB 301 TGACCTCATGCTCATCAAGCTGGCTAAGCTGCGCATGCTCAATCCCAAGTCCAGCCCT 360
QY 395 TCCCTGCGCACCAATGTCAGGCGCAGGCGCTGCTGCTACTCTCAGGTTTGAATG 454
DB 361 TCCCTGCGCACCAATGTCAGGCGCAGGCGCTGCTGCTACTCTCAGGTTTGAATG 420
QY 455 GAGCAAGAAACAGTGGCGCACACCTGACTTGGCGAGAACTGAGGCCCCCGAT 514
DB 421 GAGCAAGAAACAGTGGCGCACACCTGACTTGGCGAGAACTGAGGCCCCCGAT 480
QY 515 GTCTGATCGAGATGCCAAACAGAAACAAGAAAGCCACAGAAATCTTATGTGT 574
DB 481 GTCTGATCGAGATGCCAAACAGAAACAAGAAAGCCACAGAAATCTTATGTGT 540
QY 575 GAAATTTGTGAAATATTTAGCGCAATTTTGGGAGGTGGCCCTGCTACTGCTCACTG 634
DB 541 GAAATTTGTGAAATATTTAGCGCAATTTTGGGAGGTGGCCCTGCTACTGCTCACTG 600
QY 635 CAAGAAGCAAGCTCCAGGGAATCGAGTGGGCACTTCAATGGAGGAGGACGTCGATCTGA 694
DB 601 CAAGAAGCAAGCTCCAGGGAATCGAGTGGGCACTTCAATGGAGGAGGACGTCGATCTGA 660
QY 695 CACCAATGTTTCAAAATATGTATCTGATTTGAGAAACATGCTTAAGGACAAAGTGAGACC 754
DB 661 CACCAATGTTTCAAAATATGTATCTGATTTGAGAAACATGCTTAAGGACAAAGTGAGACC 720
QY 755 TACTTCTCCC 764
DB 721 TACTTCTCCC 730

RESULT 6
US-09-910-151-3
; Sequence 3, Application US/09910151
; Patent No. US20020137181A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 14087, A NOVEL SERINE PROTEASE MOLECULE
; FILE REFERENCE: 38155-20021.00
; CURRENT APPLICATION NUMBER: US/09/910.151
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/219,002
; NUMBER OF SEQ. ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 708
; TYPE: DNA

ORGANISM: Homo sapiens
US-09-910-151-3
Query Match 81.6%; Score 706.4; DB 10; Length 708;
Best Local Similarity 99.9%; Pred. No. 5.3e-225;
Matches 707; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 42 ATGAAATATGCTCTATTTGGGTGCTCGCTGGACATTTTCTTGCTGAATCT 101
DB 1 ATGAAATATGCTCTATTTGGGTGCTCGCTGGACATTTTCTTGCTGAATCT 60
QY 102 GTTACAGAAAGAACCTGCTCCCTATTTGGTGTACCTCAAGTCTCACTTCAACCCCTGT 161
DB 61 GTTACAGAAAGAACCTGCTCCCTATTTGGTGTACCTCAAGTCTCACTTCAACCCCTGT 120
QY 162 GTGGGCGCTCTCAATCAAAACCAAGTGGGTGTGGCCCCAGCTCACTCTATTACCAAT 221
DB 121 GTGGGCGCTCTCAATCAAAACCAAGTGGGTGTGGCCCCAGCTCACTCTATTACCAAT 180
QY 222 CTGAAAGTATGCTGGGAAATTTCAAGAGCAGATCAGAGCGGTATGAAACAAGT 281
DB 181 CTGAAAGTATGCTGGGAAATTTCAAGAGCAGATCAGAGCGGTATGAAACAAGT 240
QY 282 AACCCATTCAGATCGTCCGCTACTGGAATCAAGTCAATAGCCGCCACAGATGACCTC 341
DB 241 AACCCATTCAGATCGTCCGCTACTGGAATCAAGTCAATAGCCGCCACAGATGACCTC 300
QY 342 ATGCTCATCAAGCTGCTAAGCTGCGCATGCTCAATCCCAAGTCCAGCCCTTCCCTC 401
DB 301 ATGCTCATCAAGCTGCTAAGCTGCGCATGCTCAATCCCAAGTCCAGCCCTTCCCTC 360
QY 402 GCCACCAACCAATGTCAGGCGCAGGCGCTGCTACTCTCAGGTTTGAATGAGGCA 461
DB 361 GCCACCAACCAATGTCAGGCGCAGGCGCTGCTACTCTCAGGTTTGAATGAGGCA 420
QY 462 GAAACAGTGGCGCACACCTGACTTGGCGAGAACTGAGGCCCCCGATGCTGAT 521
DB 421 GAAACAGTGGCGCACACCTGACTTGGCGAGAACTGAGGCCCCCGATGCTGAT 480
QY 522 CGAATATGCCAAACAGAAACAAGAAAGCCACAGAAATCTTATGTGTGAATTT 581
DB 481 CGAATATGCCAAACAGAAACAAGAAAGCCACAGAAATCTTATGTGTGAATTT 540
QY 582 GTGAAATATTCAGCGCAATTTTGGGAGGTGGCCCTGCTACTGCTCACTGCAATGAC 641
DB 541 GTGAAATATTCAGCGCAATTTTGGGAGGTGGCCCTGCTACTGCTCACTGCAATGAC 600
QY 642 AACCTCCAGGGAATCGAGTGGGCACTTCAATGGAGGAGGACGTCGATCTACACCAAT 701
DB 601 AACCTCCAGGGAATCGAGTGGGCACTTCAATGGAGGAGGACGTCGATCTACACCAAT 660
QY 702 GTTTACAAATATATGCTGATTTGAGAAACATGCTTAAGGACAAAGTGA 749
DB 661 GTTTACAAATATATGCTGATTTGAGAAACATGCTTAAGGACAAAGTGA 708

RESULT 7
US-09-842-758-106
; Sequence 106, Application US/09842758
; Publication No. US20030083244A1
; GENERAL INFORMATION:
; APPLICANT: Verne, Corine A. M.
; APPLICANT: Fernandes, Bina R.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Shinkovets, Richard A.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Zernhusen, Bryan D.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Majumder, Kulud
; APPLICANT: Tchernyev, Velizar T.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patirajan, Meera

```
APPLICANT: Burgess, Catherine E
APPLICANT: Gangolli, Esna A
APPLICANT: Smithson, Glenda
APPLICANT: Raestelli, Luca
APPLICANT: MacDougall, John R
APPLICANT: Taudier, Raymond J
APPLICANT: Grose, William M
APPLICANT: Edward, Szekeres S
APPLICANT: Alabrook II, John P
TITLE OF INVENTION: No. US20030083244A1el Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 15966-783
CURRENT APPLICATION NUMBER: US/09/842,758
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: 60/200,158
PRIOR FILING DATE: 2000-04-26
PRIOR APPLICATION NUMBER: 60/200,613
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,780
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/201,006
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,007
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,236
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,238
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,186
PRIOR FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: 60/201,474
PRIOR FILING DATE: 2000-05-03
PRIOR APPLICATION NUMBER: 60/201,508
PRIOR FILING DATE: 2000-05-03
PRIOR APPLICATION NUMBER: 60/220,591
PRIOR FILING DATE: 2000-07-25
PRIOR APPLICATION NUMBER: 60/232,678
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 60/263,217
PRIOR FILING DATE: 2001-01-22
PRIOR APPLICATION NUMBER: 60/265,160
PRIOR FILING DATE: 2001-01-30
NUMBER OF SEQ ID NOS: 113
SOFTWARE: PatentIn Ver. 2.1
LENGTH: 721
TYPE: DNA
ORGANISM: Homo sapiens
US-09-842-758-106

Query Match
Best Local Similarity 81.3%; Score 703.8; DB 11; Length 721;
Matches 716; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
```

```
QY 339 CTCATGCTCATCAAGCTGCTAAGCCCTGCATGCTCAATCCCAAGTCCAGCCCTTCCC 398
DB 302 CTCATGCTCATCAAGCTGCTAAGCCCTGCATGCTCAATCCCAAGTCCAGCCCTTACC 361
QY 399 CTCGCCACCAACATGTCAGGCCAGGCACTGCTCTATCTTCAAGTTTGGACTGGAGC 458
DB 362 CTCGCCACCAACATGTCAGGCCAGGCACTGCTCTATCTTCAAGTTTGGACTGGAGC 421
QY 459 CAAGAAAACAGTGGCCGACACCCCTGACTTGGGAGAACCTGGAGGCCCGGTGATGCT 518
DB 422 CAAGAAAACAGTGGCCGACACCCCTGACTTGGGAGAACCTGGAGGCCCGGTGATGCT 481
QY 519 GATGAGAAATGCC-AAAAACAGAACAGAAAAAGCCACAGAAATCTTATGTGAA 577
DB 482 GATGAGAAATGCCAAAAACAGAACAGAAAAAGCCACAGAAATCTTATGTGAA 541
QY 578 ATTGTGAAAGTATTCAGCCGAATTTTGGGAGGTGGCCGTGCTACTGCTCAATCTGCA 637
DB 542 ATTGTGAAAGTATTCAGCCGAATTTTGGGAGGTGGCCGTGCTACTGCTCAATCTGCA 601
QY 638 AGACAGCTCCAGGGAATTCAGGTGGGCACTTCATGAGAGGAGCGTGGCATCTACAC 697
DB 602 AGACAGCTCCAGGGAATTCAGGTGGGCACTTCATGAGAGGAGCGTGGCATCTACAC 661
QY 698 CAATGTTTACAAATATGATCTGATTTGAGAACACTGCTAAGGACAAAGTGAACCTTA 756
DB 662 CAATGTTTACAAATATGATCTGATTTGAGAACACTGCTAAGGACAAAGTGAACCTTA 720
```

RESULT 8
US-09-888-615-43
Sequence 43, Application US/09888615
Patent No. US20020064856A1

GENERAL INFORMATION:

APPLICANT: PLOMAN, GREGORY
APPLICANT: WHITE, DAVID
APPLICANT: CAENEPEL, SEAN
APPLICANT: CHARYDCZAK, GLEN
APPLICANT: MANNING, GERARD
APPLICANT: SUDARSANAM, SUCHA
TITLE OF INVENTION: NOVEL PROTEASES

FILE REFERENCE: 038602/1214
CURRENT APPLICATION NUMBER: US/09/888,615
CURRENT FILING DATE: 2001-06-26
PRIOR APPLICATION NUMBER: 60/214,047
PRIOR FILING DATE: 2000-06-26
NUMBER OF SEQ ID NOS: 150
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 43
LENGTH: 798
TYPE: DNA
ORGANISM: Homo sapiens
US-09-888-615-43

Query Match
Best Local Similarity 70.0%; Score 606.4; DB 9; Length 798;
Matches 707; Conservative 0; Mismatches 1; Indels 90; Gaps 1;

```
QY 42 ATGAATATATGCTTCTATTTGGGTGCTCGCTGGACATTTTCTTGGACTCATCT 101
DB 1 ATGAATATATGCTTCTATTTGGGTGCTCGCTGGACATTTTCTTGGACTCATCT 60
QY 102 GTTCAGAAAGAACCCGCTCCCTATTTGGTGTACCTCAAGTCACTTCAACCCCTGT 161
DB 61 GTTCAGAAAGAACCCGCTCCCTATTTGGTGTACCTCAAGTCACTTCAACCCCTGT 120
QY 162 GTGGCGTCTCATCAACCCAGCTGGGTGCTGGCCCGACGCTCACTGCTATTTCAAAT 221
DB 121 GTGGCGTCTCATCAACCCAGCTGGGTGCTGGCCCGACGCTCACTGCTATTTCAAAT 180
QY 222 CTGAAGATGATGCTGGGAAATTTCAAGAGCAGAGTCAAGACGGTACTGAACAGCAAT 281
DB 181 CTGAAGATGATGCTGGGAAATTTCAAGAGCAGAGTCAAGACGGTACTGAACAGCAAT 240
```

Oy	282	AAACCCATTCAATGTCGCGTACTGGAATACAGTCAATAGGCCCCCAAGATGACCTC	3411
Oy	282	AAACCCATTCAATGTCGCGTACTGGAATACAGTCAATAGGCCCCCAAGATGACCTC	3411
Db	241	AAACCCATTCAATGTCGCGTACTGGAATACAGTCAATAGGCCCCCAAGATGACCTC	300
Oy	342	ATGCTCATCAAGCTGGCTTAAGCTCGCATGCTCAATCCGAAAGTCAGCCCTTCCCTC	401
Db	301	ATGCTCATCAAGCTGGCTTAAGCTCGCATGCTCAATCCGAAAGTCAGCCCTTACCTC	360
Oy	402	GCCACCAACCAATGTCAAGCCAGGCACTGTCTGTCTACTCTCAGGTTTGGACTGAGCCAA	461
Db	361	GCCACCAACCAATGTCAAGCCAGGCACTGTCTGTCTACTCTCAGGTTTGGACTGAGCCAA	420
Oy	462	GAAAAACGT----- 470	
Db	421	GAAAAACGTGGCTTTGGCACTGAGCCAGCCACAGGCCATCTACTCTGCACAGAGCCCA	480
Oy	471	-----GCGGCACACCTGACTTGGCG	491
Db	481	GCCATTCCTGATTGGCAGAGACCAATTCATGAAACAAGCGGCAGACCTTGACTTGGCG	540
Oy	492	CAGAACCTTGAGAGCCCCCTGATGTGTGATTCGAGAAATGCCAAAAACAGAACAGAAAA	551
Db	541	CAGAACCTTGAGAGCCCCCTGATGTGTGATTCGAGAAATGCCAAAAACAGAACAGAAAA	600
Oy	552	AGCCACAGGAATTCCTTATGTGTGAAATTGTGAAAGTATTGACCGGAATTTTGGGGAG	611
Db	601	AGCCACAGGAATTCCTTATGTGTGAAATTGTGAAAGTATTGACCGGAATTTTGGGGAG	660
Oy	612	GTGGCGGTGCTACTGTCTCTCTGCAAGACAAGCTCCAGGGAAATCGAGGTGGGGCACTTC	671
Db	661	GTGGCGGTGCTACTGTCTCTCTGCAAGACAAGCTCCAGGGAAATCGAGGTGGGGCACTTC	720
Oy	672	ATGGAGGGGAGCTGCGGCACTTACACCAATGTTTAAATATGTATCTTGAGATTGGAAC	731
Db	721	ATGGAGGGGAGCTGCGGCACTTACACCAATGTTTAAATATGTATCTTGAGATTGGAAC	780
Oy	732	ACTGCTAAGACCAAGTGA 749	
Db	781	ACTGCTAAGACCAAGTGA 798	
RESULT 9			
US-10-246-006-3			
: Sequence 3, Application US/10246006			
: Publication No. US2003014493A1			
: GENERAL INFORMATION:			
: APPLICANT: Conklin, Darrell C.			
: TITLE OF INVENTION: Human Serine Protease			
: FILE REFERENCE: 00-16			
: CURRENT APPLICATION NUMBER: US/10/246,006			
: CURRENT FILING DATE: 2002-09-17			
: NUMBER OF SEQ ID NOS: 4			
: SOFTWARE: FastSeq for Windows Version 3.0			
: SEQ ID NO 3			
: LENGTH: 705			
: TYPE: DNA			
: ORGANISM: Artificial Sequence			
: FEATURE:			
: OTHER INFORMATION: This degenerate nucleotide sequence encodes the			
: OTHER INFORMATION: amino acid sequence of SEQ ID NO:2.			
: FEATURE:			
: NAME/KEY: misc feature			
: LOCATION: (1) _ (705)			
: OTHER INFORMATION: n = A,T,C or G			
: US-10-246-006-3			

[illegible]

Db	1	ATGAARTAYGNTTAYTAYTNGAGNTNTYNGCNGAGNACNTTAYTAYTAYTGCGNAYWMSMNS	60
Qy	102	GTTCAGAAAGAACGCCCTGCTCCCAATTTGGTGTCACCTAACCTCAACCCCTGT	161
Db	61	GTNCAPARARGARGAACCGCNCNCNTAYTNGTNTAYTNTAAWMSNCAYTTAAVCCNTG	120
Qy	162	GTGGGCGCTCCTCATCAAAACCCAGCTGGGTGCTGGCCCCAGCTCACTGCTATTYTAACCAAT	221
Db	121	GTNGGNGTNTYNTATTPARCCNMSNTGGGTNTYNTGNCNCNGCNCAYTGYTAYTNGCMAAY	180
Qy	222	CTGAAAGTGATGCTGGGAAATTTTCAAGAGCAGAGTCAGAGACGGTACTGTAAACAGCAATT	281
Db	181	YTNAARGTNAAGYTNAGTNGAAAYTTAYARSMNGMNGTNGMGAAGGAGNACNGARCAPACNAH	240
Qy	282	AAACCCATTCAGATCGTCCGCTACTGTGAAACTACAGTCAATGAGGGCCCCACAGATGACCTC	341
Db	241	AAAYCCNATHCARATHGTMTGNTAYTGGAAAYTAYWNCAYTWSGNCNCNCARBAHYATYN	300
Qy	342	ATGCTCATCAAGCTGGCTAAGCTGAGCTCCACTGCTCAATCCCAAGTCCAGCCCTTCCCTC	401
Db	301	ATGYTNAATHAARYTNGCNAARCCNGCNAATGYTNAAYCCNAARGTNCARCCNTYNA	360
Qy	402	GCCACCAACCAATGTCAGGCGCAGGCACTGTCGTCACTCACTGAGTTTGACCTGGAGCCAA	461
Db	361	GCNACNACNAAYGTNMGCNCNGNACNGTNTGYTNTYTWWSGNYTNGAYTGGWSNCAR	420
Qy	462	GAAGAACAGTGGCCGACACCTGACTTGGCGGAGAACTGGAGGCCCCCGCTGATGTCAT	521
Db	421	GABAAWYNGNGMGCAACNGAAYTNTNGNCARBAAYTNTGARGCNCNGTNAATGWSNGAY	480
Qy	522	CGAGAATGCCAAAAAACAAGCAAGGAAAAAGCCACAGAAATCTCTATGTCGAAATTT	581
Db	481	WNGNARTGYCARAABAACNGARCARGANAARMSNCAWNGAALWNSYTNNTGYTNAARTTY	540
Qy	582	GTGAAGATATTCAAGCCGAATTTTGGGGAGGTGGCCGTGTGCTACTGTCATCTGCAAGAAC	641
Db	541	GTNAARGTNTTYYWMSNGNAHTTYYGNGARGTNGCNGTNGCNAACGNTNAHTHYAARGA	600
Qy	642	AAAGCTCCAGGAAATCGAGGTGGGGCACTTCATGGAAGGGGACGTGGGCATCTACACCAAT	701
Db	601	AARYTNCARGNATHGARGTNGNCAAYTTAYTGGGNGNGAAGTNGGNAHTHAYACNAAY	660
Qy	702	GTTTACAAATATGTATCTCGATTGAGAAACCTGTTAAGACAAG	746
Db	661	GTNTAYAAARTAYGTNWSMTGATHTGARBAAYACNGCMAARGAYAAR	705

```

RESULT 10
US-10-101-510-199
: Sequence 199. Application US/10101510
: Publication No. US20030148295A1
: GENERAL INFORMATION:
: APPLICANT: WAN, JACKSON
: APPLICANT: WANG, YIXIN
: TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
: FILE REFERENCE: 15117, 0012
: CURRENT APPLICATION NUMBER: US/10/101,510
: CURRENT FILING DATE: 2002-03-20
: PRIOR APPLICATION NUMBER: 60/276,947
: PRIOR FILING DATE: 2001-03-20
: NUMBER OF SEQ ID NOS: 805
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 199
: LENGTH: 853
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-101-510-199

```

Query	109	AAGAAGACCCGCTCCCTATTGCGTACCAAGTCACTTCAACCCCTGTGTGGCG	168
Query Match	17.8%	Score 154.4; DB 12; Length 853;	
Best Local Similarity	55.0%	Pred. No. 1.9e+40;	
Matches 349; Conservative	0;	Mismatches 276; Indels 9; Gaps 2.	

Db 135 AGAGAAATTCCTCCCTACAGAGGTGCTGAAATTCGCTCCCACTTGCAGTGGCT 194
Qy 169 TCTCATCAAAACCAAGCTGGGTCTGGCCCAAGCTCACTGTATTTACCAATCTGAAG 228
Db 195 CCTCATCAGGAAACAGTGGGTGTATCAGACGCTCATGTGTACAAAGCCCGCATCCAG 254
Qy 229 TGATGCTGGGAAATTTCAAGAGCAGAGTCAGAGACGGTACTGAACAGACAAATTAACCCA 288
Db 255 TGAAGCTGGAGAGACACAACTCAAAATCTGGAGGGAAATGAGCATGTTCAATGCGG 314
Qy 289 TTCAGATGCTCCGCTACTGAACTACAGTCAATAGGCCCCAAGAGTACCTCATGCTCA 348
Db 315 CCAAGATCATCCGACACCTTAATTAACAGAGACACTGTGACAAATGACATCATGCTGA 374
Qy 349 TCAAGCTGGCTAAGCTGCTCAATGCTCAATCCCAAGTCCAGCCCTTCCCTCCGACCA 408
Db 375 TCAAACTCTCTCACTGCTCCGTCATCAATGCCCCGCTGTCCACATCTCTGCCCCACG 434
Qy 409 CCAATGTGAGGCAAGGCACTGTCTGTACTCTCAAGTTTGGACTGAGCCAGAAACA 468
Db 435 CCCCTCAGCTGCTGGCACTGAGTCTCTCATCTCCGCTGGGCAACCTTGAGCTTG 494
Qy 469 GTGGCCGACACCCCTGACTTGGGCGCAGAACTGAGGCCCCGCTGATGTGATCGAGAT 528
Db 495 GTGCTGACTACCCAGACAGAGCTGAAGTCTGGATGCTCCGGTGTGACCCAGGCTGAGT 554
Qy 529 GCCAAAAACAGAACAGAAAAACCAAGGAATCTCTTATGTGTAAATTTGTGAAG 588
Db 555 GTAAAGCTCTTACCTGGAAGAAATTAACACAGCATGTTCTGTGTGGCTTCTTGAGG 614
Qy 589 TATTCAGCCGAATTTTGGGAGGTGGCGCTGTACTGTCACTGCAAGCAAGAGCTCC 648
Db 615 GAGGAGAGATTTCTCCAGAGGTGACTGTGGTCCCTGTGTGCTGCAACGAGAGCTCC 674
Qy 649 AGGGAATCG--AGTGGGGCACTTTCATGAGGAGGAGCTC-----GGCATCTACACCA 699
Db 675 AAGAGTGTCTCCCTGGGGGCATGCTGTGCTGGAAGAACAGGCTGAGTCAACA 734
Qy 700 ATGTTACAAATATGTATCTGTGATTTGAACAC 733
Db 735 AGTCTACAACTATGTGAGCTGATTAAGGACAC 768

RESULT 11
US-09-954-456-748
Sequence 748, Application US/09954456
Patent No. US20020115057A1
GENERAL INFORMATION:
APPLICANT: Young, Paul
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can
FILE REFERENCE: 689290-76
CURRENT APPLICATION NUMBER: US/09/954,456
CURRENT FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/60/233,617
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,052
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,923
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,134
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,637
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,638
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,711
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,720
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,840
PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: US/60/235,863
PRIOR FILING DATE: 2000-09-27
NUMBER OF SEQ ID NOS: 2276
SOFTWARE: PatentIn version 3.0
SEQ ID NO 748
LENGTH: 850
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-456-748

Query Match 17.8%; Score 154; DB 10; Length 850;
Best Local Similarity 55.1%; Pred. No. 2.6e-40;
Matches 348; Conservative 0; Mismatches 275; Indels 9; Gaps 2;

Qy 111 GAAGACCTGCTCCCTTATTTGTTGTTACTCAAGTCTCACTTCAACCCCTGTGTGGGCTC 170
Db 134 GAGAAATCTCTCCCTACAGAGTGTCTGTAATTTGTGGCTCCCACTTCTGGGGTGTCC 193
Qy 171 CTCATCAACCCAGCTGGGTGCTGGCCCAAGCTCACTGTATTTAACCAATCTGAAGTG 230
Db 194 CTCATCAAGCGAACAGTGGGTGTATCAGACGCTCACTGCTACAAAGACCCCATCAGGTG 253
Qy 231 ATGCTGGAAATTTCAAGACAGAGTCAGAGCGTACTGAACAGACAAATTAACCCATT 290
Db 254 AGACTGGAGAGCAACATCAAGTCTGTGAGGGGAATAGAGCTTCAATCAATGCGGCC 313
Qy 291 CAGATGCTCCGCTACTGGAATCACTAGTACGAGCCCAAGATGACCTCATGCTCATC 350
Db 314 AAGATCAATCGCACCTTAATTAACAGAGACACTGTGACAAATGACATCATCTGATC 373
Qy 351 AAGCTGCTAAGCTGTCATGCTCAATCCCAAGTCCAGCCCTTCCCTGCGACAC 410
Db 374 AAATCTCTCACTGCGCTGATCAATGCCGCTGTCACATCTCTGCGCACCGCC 433
Qy 411 AATGTACAGCCAGCACTGTCTGTACTCTCAAGTTTGAAGTGAAGCAAGAAACAGT 470
Db 434 CCTCAGCTGTGCACTGAGTGTCTCATCTCCGCTGGGCAACACTGTGAGCTTTGT 493
Qy 471 GCGCGACACCTGACTTGGCGCAGAACCTGAGGCCCCCGTGTGATGTGAGAAATGC 530
Db 494 GCTGACTACCCAGACAGCTGAAGTGTCTGTGATCTCGGTGCTGACCCAGGCTGAGTGT 553
Qy 531 CAAAAACAGAACAGAAAAAGCCACAGAAATCTTATGTGTGAATTTGTGAAGTA 590
Db 554 AAGGCTCTTACCTCGGAAGAAATTAACACAGCATGTTGTGTGGCTTCTTGAGGGA 613
Qy 591 TTCAAGCGAATTTTGGGAGGTGGCCGTTCTACTGTCACTGTGCAAGACACTCCAG 650
Db 614 GCGAAGGATTCCTCCAGCGTCACTGTGTGGCCCTGTGTCTCAACGAGACACTCCAA 673
Qy 651 GGAATCG--AGTGGGGCACTTCATGGAAGGAGGAGCTC-----GGCATCTACACCAAT 701
Db 674 GAGTGTCTCTCGGGGCATGCTGTGCTGGAAGAACAGGCTGAGTCTACACCAAG 733
Qy 702 GTTACAAATATGTATCTGTGATTTGAACAC 733
Db 734 GTCTACAACTATGTGACTGATTAAGGACAC 765

RESULT 12
US-09-923-779-148
Sequence 148, Application US/09923779
Patent No. US20020076721A1
GENERAL INFORMATION:
APPLICANT: Pyle, Ruth A.
APPLICANT: Xu, Jiangchun
APPLICANT: Kalos, Michael D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.553
CURRENT APPLICATION NUMBER: US/09/923,779
CURRENT FILING DATE: 2001-08-06
NUMBER OF SEQ ID NOS: 155


```

; SOFTWARE: FastSeq for Windows Version 4.0.0
; SEQ ID NO: 148
; LENGTH: 802
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-922-779-148

```

Query Match	17.7%;	Score 153.2;	DB 9;	Length 802;
Best Local Similarity	53.9%;	Pred. No. 4.6e-40;		
Matches 342; Conservative	0;	Mismatches 283;	Indels 9;	Gaps 1

OY	109	AAGAAGACCCCTGCTCCCTATTTGGTATACCTCAAGTCACTCACTCAACCCCTGATGAGCG	168
Db	98	AGGAGATTCTGTGCCCCCTACACAGGTGTCCTTGAATTTCTGGCTACCACTTTCTGGGTGGCT	157
OY	169	TCCTCATCAAAACCCAGCTGGGTGCTGGCCCCAGCTCACTGCTATTTTACCAAAATCTGAAG	228
Db	158	CCCTCATTCGCGAACAGTGGGTGGTGTCAAGCAGATCTCTCTCAAGTCCCGCATCCAGG	217
OY	229	TGATGTGGGAAATTTCAAGACAGAGTCAAGAACGGTACTGTAACAGACAAATTAAACCCA	288
Db	218	TGAGACTGGAGAGCAACACATGAAAGTCTTGAGGGGAAATGAACAATTATCATATCGG	277
OY	289	TTCAAGATCGTCCGCTACTGGAACTACAGTCAATAGCGCCCCACAGATCACTCATGCTCA	348
Db	278	CCAAAGTATATCGCGCACCCCAATATACACAGCGGCACTTGGACAATGACATCTGTGA	337
OY	349	TCAAGTGGCTAAAGCTTCGCTCATCTCAATCCCAAATGTCAGAGCCCTTCCCTGCGCACCA	408
Db	338	TCMACTCTCTCACTCCGCGGTATCAATATTTCCGCGGTGCGGCATCTCTGTGCCACTG	397
OY	409	CCAAATGTCAGGCGAGGCACTGTCTGTCTACTCTCAGTTTGGACTGGAGCGCAAGAAAACA	468
Db	398	CCCCCTCAGCTGTGGCACCGAATCCCTCATCTTCGGCTGGGGCAACACTCTGAGTTCTG	457
OY	469	GTGGCGGACACCTCTGACTTTCGGGAGAACTTGAAGGCCCGGTGATGTCTGATCGAGAT	528
Db	458	GTGCGGACTACCCAGACGAGCTGCAAGTGCCTGATGTCTCTGTGCTGAGCGAAGCTAGT	517
OY	529	GCCAAAAAACAAACAAGAAAAAGCCACAGAAATCTTATGTGTGAATTTGTGAAG	588
Db	518	GTGAAGCTCTCTACCTTGAAAGATTACCAACAACATGTTCTGTGTGGGCTTCTCTGAGG	577
OY	589	TATTCAGCCGAATTTTGGGAGGTGAGCGCTGTCTACTGTCACTGTGCAAGAAGCAAGCTCC	648
Db	578	GAGGCAAGGATTCGTGCCAGGGGTGATTTCTGTGTGGCCCTGTGTGTTCCAAATGAGAGCTCC	637
OY	649	AGGGAATCGAGGTGGGGCACTTCAATGGGAGGGGACGTGGCA-----TCTAACCA	699
Db	638	AAGGAATGTCTCTGTGGGCTATAGCTGTGCCAGAAAGACAGGCTGAGTCTACACCA	697
OY	700	ATGTTTAAACAATATGATTCCTGGAATGAGAACAC	733
Db	698	AGGTCTACACTAATGTGACATGGATTTAAGGACAC	731

```

RESULT 13
US-10-101-510-210
: Sequence 210, Application US/10101510
: Publication No. US20030148295A1
: GENERAL INFORMATION:
: APPLICANT: WAN, JACKSON
: APPLICANT: WANG, YIXIN
: TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
: FILE REFERENCE: 15117.0012
: CURRENT APPLICATION NUMBER: US/10/101,510
: CURRENT FILING DATE: 2002-03-20
: PRIOR APPLICATION NUMBER: 60/276,947
: PRIOR FILING DATE: 2001-03-20
: NUMBER OF SEQ ID NOS: 805
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 210
:
: LENGTH: 802

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-210

```

Query Match	17.7%	Score 153.2	DB 12	Length 802
Best Local Similarity	53.9%	Prod. No. 4.6e-40		
Matches 342	Conservative 0	Mismatches 283	Indels 9	Gaps 1

Oy	109	AAGAAAGACCCGTGCTCCCTAATTTGGGTGACCTCAAGTCTCACTTCAACACCCCTGTGGGCG	168
Db	98	AGGAGAAATTCTGTCCCTTACCAAGTGTCTTGAATTTCTGGCTAACCACTTCTGGGGTGCT	157
Oy	169	TCTCTATCAAAACCCAGCTGGGTGCTGGCCCCAGCTCACTGTCTATTATCCAAATCTGAAG	228
Db	158	CCCTCATAGGAAACAGTGGGGTGTGACAGGTCTACGTCTAACAGTCCCGCATCAGG	217
Oy	229	TGATGCTGGGAAATTTCAAGACAGACAGTACAGAGGTCTGAAACAGCAATTAACCCCA	288
Db	218	TGAGACTGGGAGACAAACATCGAAGTCTTGGAGGGAGATAAACGTTTCAATCAATGGG	277
Oy	289	TTCAAGTGTCCGCTACTGTGAACCTACAGTCAATAGCGCCCCACAGATGACCTTATGCTCA	348
Db	278	CCAAGATCATCCGCAACCCCAAAATACAAACGCCGAGCTCTGACAAATGACATCTGTGTA	337
Oy	349	TCAAGCTGGCTAAGCTTGTCCATGTCTCAATCCCAAGTCCAGCCCCCTTCCCCCTGCCACCA	408
Db	338	TCAAGCTCTCTCACTGCTCCGTCATCAATTTCCCGGTGTCGCGCANTCTCTGCCCCACTG	397
Oy	409	CCAATGTACAGGCCAGGACACTGTCTCTACTCTCAGGTGTGGACTGAGCCCAAGAAACA	468
Db	398	CCCTTCCAGCTGTGTGGCACCCAGTCCCTCATTTCCGCGTGGGGCAACACTGTGATTCTG	457
Oy	469	GTGGCCGACACCTTGACTTTCGGGCAAAACTTGAAGCCCCCGTATGTCTGATCGAAT	528
Db	458	GTGCGCACTACCCAGACAGAGCTGCAGTGCCTGAGTGTCTGTGTCGACCAAGCTGAGT	517
Oy	529	GCCAAAAACAGAACAGAAAAAGCCACAGGAATTCCTATGTGGAATTTGTAAAG	588
Db	518	GTGAAGCTCTTACCTCTGAAAAGATTACCAACAACTGTTCTGTGTGGGCTTCTTGAGG	577
Oy	589	TATTCAGCCGAATTTTGGGAGGTGGCCGTGCTACTGTCACTTGCACAAAGACAAGCTCC	648
Db	578	GAGGCAAGGAATTCGTGCCAGGGGTGATTTCTGGTGGCCCTGTGTCTTCATATGAGAGCTCC	637
Oy	649	AGGGAATCGAGGTGGGCACTTTCATGGAGGGGACGTCCGCA-----TCTACACCA	699
Db	638	AAGGAATGTCTCTCGGGGCTATGGCTGTGCCACAAGAACAGGCGCTGGAGTCTACACCA	697
Oy	700	ATGTTTCAAAATATGTATCTCGATGTGAAACAC	733
Db	698	AGGCTTCAACTATGTGACTGTGAATTAAGACAC	731

RESULT 14
US-10-181-808-1
Sequence 1, Application US/10181808
Publication No. US20030157634A1
GENERAL INFORMATION:
APPLICANT: Polymun Scientific Immunobiologische Forschung GmbH
APPLICANT: Matlanovich, Diethard
APPLICANT: Kallinger, Hermann
APPLICANT: Hohenblum, Hubertus
APPLICANT: Naschberger, Stefan
APPLICANT: Weik, Robert
TITLE OF INVENTION: METHOD FOR THE MANUFACTURE OF RECOMBINANT TRYPSIN
FILE REFERENCE: 2235-146
CURRENT APPLICATION NUMBER: US/10/181,808
CURRENT FILING DATE: 2002-07-23
PRIOR APPLICATION NUMBER: PCT/EP01/00770
PRIOR FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: US 60/177,348
PRIOR FILING DATE: 2000-01-24
NUMBER OF SEQ ID NOS: 8

SOFTWARE: PatentIn version 3.1
 SEQ ID NO 1
 LENGTH: 699
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-181-808-1

Query Match 16.4%; Score 142; DB 12; Length 699;
 Best Local Similarity 52.8%; Pred. No. 2.3e-35;
 Matches 335; Conservative 0; Mismatches 290; Indels 9; Gaps 1;

109 AAGAAGACCTGCTCCCTATTGGTGTACCTCAAGTCTCACTTCAACCCCTGTGTGGCG 168
 47 AGGAGATTCTGTCCCTACAGGTGTCCCTGAATTTGGCTACATTTCTGTGTGGCT 106
 169 TCCTCATCAAAACCAAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 228
 107 CCTCATCAAAACCAAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 166
 229 TGATGTGGGAAATTTCAAGACAGTCAAGACGCTACTGAAACAGCAATTAAACCCA 288
 167 TGAGACTGGAGAGACAAACATCAAGTCTGTGGAGGGGAAATGACAGTTCAATCAG 226
 289 TTCAAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 348
 227 CCAAGATCATCTCCGACCCCAATGACAGAGAACTCTGAAACATGACATCTTTAA 286
 349 TCAAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 408
 287 TCAAGTGTGTCTCTCAAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 346
 409 CCAATGTCAAGGCGACGACTGTCTGTCTACTCTCAAGTTTGGATGTGAACCAAGAAACA 468
 347 CCCCTCAGCAGCTGGGACGAGAGTGCCTCACTCTGGGTGGGCAACATGCGAGCTGTG 406
 469 GTGGCCGACACCTGACTTGGGAGGAACTGGAGGCGCCGCTGATGTGTAGAGAT 528
 407 GCGCGGACTACCGAGAGAGTGAAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 466
 529 GCCAAATAACAGAACAGAAAGAAAGCCACAGAAATTCCTTATGTGTGAATTTGTGAAG 588
 467 GTGAAGCTCTTACCTCGGAAAGATTAACAGCAATGTTCTGTGTGGGCTTCTGTAGG 526
 589 TATTCAGCCGAATTTTGGGAGGTGGCGTGTCTACTGTCAATGTCAAAAGCAAGCTCC 648
 527 GAGGCAAGATTCATGTCAAGGTGATTTCTGTGGCCCTGTGTGTCAATGAGACAGCTCC 586
 649 AGGGAATCGAG-----GTGGGCACTTCATGGGAGGGGAGCGTGGCATTTACACCA 699
 587 AAGAGATTGTCTCTGGGGTATGTGCTGTGCCAGAGAACAGAGCTGAGTTACACCA 646
 700 ATGTTTCAATATATATCTGTGATTTGAGAACAC 733
 647 AGGTCTACACTACGTGAATGATTAAGAACAC 680

RESULT 15

US-09-923-779-91
 Sequence 91, Application US/09923779
 Patent No. US20020076721A1
 GENERAL INFORMATION:
 APPLICANT: Pyle, Ruth A.
 APPLICANT: Xu, Jiangchun
 APPLICANT: Kalos, Michael D.
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
 FILE REFERENCE: 210121.553
 CURRENT APPLICATION NUMBER: US/09/923,779
 CURRENT FILING DATE: 2001-08-06
 NUMBER OF SEQ ID NOS: 155
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 91
 LENGTH: 522

TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 3, 5
 OTHER INFORMATION: n = A,T,C or G
 US-09-923-779-91

Query Match 15.3%; Score 132.8; DB 9; Length 522;
 Best Local Similarity 55.1%; Pred. No. 2.3e-33;
 Matches 260; Conservative 0; Mismatches 212; Indels 0; Gaps 0;

109 AAGAAGACCTGCTCCCTATTGGTGTACCTCAAGTCTCACTTCAACCCCTGTGTGGCG 168
 50 AGGAGATTCTGTCCCTACAGGTGTCCCTGAATTTGGCTACATTTCTGTGTGGCT 109
 169 TCCTCATCAAAACCAAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 228
 110 CCTCATCAAGCAAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 169
 229 TGATGTGGGAAATTTCAAGACAGTCAAGACGCTACTGAAACAGCAATTAAACCCA 288
 170 TGAGACTGGAGAGACAAACATCAAGTCTGTGGAGGGGAAATGACAGTTCAATCAG 229
 289 TTCAAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 348
 230 CCAAGATCATCTCCGACCCCAATGACAGAGAACTCTGAAACATGACATCTTTAA 289
 349 TCAAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 408
 290 TCAAGTGTCTCTCAAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 349
 409 CCAATGTCAAGGCGACGACTGTCTGTCTACTCTCAAGTTTGGATGTGAACCAAGAAACA 468
 350 CCCCTCAGCAGCTGGGACGAGTGCCTCACTCTGGGTGGGCAACATGCGAGCTGTG 409
 469 GTGGCCGACACCTGACTTGGGAGGAACTGGAGGCGCCGCTGATGTGTAGAGAT 528
 410 GTGCCGACTACCGAGAGAGTGAAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 469
 529 GCCAAATAACAGAACAGAAAGAAAGCCACAGAAATTCCTTATGTGTGAATTTGTGAAG 580
 470 GTGAAGCTCTTACCTCGGAAAGATTAACAGCAATGTTCTGTGTGGGCTT 521

Search completed: November 8, 2003, 00:24:39
 Job time : 357 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 8, 2003, 01:23:17 ; Search time 29 Seconds
(without alignments)
5743.592 Million cell updates/sec

Title: US-10-037-270-482
Perfect score: 1560
Sequence: 1 ggcactactccctgcagcta.....tcggaaatgcagcactagc 866

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 566616

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p,model -DEV=xjh
-Q=/cgn2_1/USPTO.epool/US10037270/runat_07112003_140516_8805/app.query.fasta_1.1031
-DB=PIR 76 -QMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomum62 -TRANS=humam40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USFR=US10037270.@CGN_1.1.25@runat_07112003_140516_8805 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	377.5	24.2	248	2	SS5067
2	353	22.6	304	2	S33496
3	352.5	22.6	247	1	A25852
4	351.5	22.5	246	2	B25528
5	350.5	22.5	246	1	TRDGC
6	350.5	22.5	259	2	I38363
7	349	22.4	243	2	A35871
8	346.5	22.2	246	1	TRRT1
9	346	22.2	247	2	TRDG
10	346	22.2	247	2	S12764
11	344.5	22.1	248	2	SS5066
12	340.5	21.8	246	1	TRRT2
13	335.5	21.5	231	1	TRPGTR
14	334.5	21.4	247	1	B25852
					trypsin (EC 3.4.21

15	329.5	21.1	247	2	S13813	trypsin (EC 3.4.21
16	328.5	21.1	247	2	A27547	trypsin (EC 3.4.21
17	312.5	20.0	229	1	TRB0TR	trypsin (EC 3.4.21
18	312	20.0	229	1	TRDS	trypsin (EC 3.4.21
19	312	20.0	247	2	S05494	trypsin (EC 3.4.21
20	311.5	20.0	228	2	S31779	trypsin (EC 3.4.21
21	311	19.9	246	2	Q01472	trypsin (EC 3.4.21
22	309	19.8	246	2	Q01471	trypsin (EC 3.4.21
23	287	18.4	242	2	S49489	trypsin (EC 3.4.21
24	285	18.3	242	2	S31778	trypsin (EC 3.4.21
25	280	17.9	242	2	S31775	trypsin (EC 3.4.21
26	280	17.9	242	2	S31776	trypsin (EC 3.4.21
27	264.5	17.0	240	2	S39047	trypsin (EC 3.4.21
28	260	16.7	241	2	S39048	trypsin (EC 3.4.21
29	247.5	15.9	255	2	A27122	cathepsin G (EC 3.
30	244.5	15.7	250	2	T01779	trypsin (EC 3.4.21
31	244	15.6	248	2	S33755	granzyme-like prot
32	243	15.6	260	2	S15659	granzyme-like prot
33	237	15.2	248	2	S43259	cathepsin G (EC 3.
34	235.5	15.1	261	2	S40162	granzyme C (EC 3.4
35	235	15.1	248	1	PRM5C2	natural killer cel
36	231	14.8	248	2	A43520	venomolysin A (EC 3.4
37	228.5	14.6	233	1	UG0169	granzyme B (EC 3.4
38	228.5	14.6	247	1	PRM5CL	7S nerve growth fa
39	221	14.2	261	1	NGMSG	plasmin (EC 3.4.21
40	216.5	13.9	812	1	PLBO	cytotoxic T-lympho
41	215	13.8	246	2	A32692	serine proteinase
42	214	13.7	253	2	A53968	tissue kallikrein
43	214	13.7	261	2	A41020	complement factor
44	209.5	13.4	263	1	I55608	trypsin (EC 3.4.21
45	208.5	13.4	250	2	S31384	

ALIGNMENTS

RESULT 1
SS5067
trypsin (EC 3.4.21.4) I precursor, pancreatic - chicken
N/Alternate names: trypsinogen I
C/Species: Gallus gallus (chicken)
C/Date: 28-Oct-1996 #sequence revision 07-Feb-1997 #text_change 21-Jul-2000
C/Accession: SS5067; S72345; SS5065; S72346; S71155
R/Mang, K.; Gan, L.; Lee, I.; Hood, L.
Biochem. J. 307, 471-479, 1995
A/Title: Isolation and characterization of the chicken trypsinogen gene family.
A/Reference number: SS5065; MUID:95251611; PMID:7733885
A/Accession: SS5067
A/Molecule type: mRNA
A/Residues: 1-248 <MAN1>
A/Cross-references: EMBL:U15156; NID:G603904
A/Experimental source: clone 1-P38
A/Accession: S72345
A/Molecule type: DNA
A/Residues: 1-248 <MAN2>
A/Cross-references: EMBL:U15155; NID:G603902; PIDN:AAA79912.1; PID:G603903
A/Experimental source: clone 1-P38
A/Accession: S72346
A/Molecule type: DNA
A/Residues: 1-9, 'V', '11-12', 'T', '14-102', 'A', '104-214', 'I', '216-248 <MAN4>
A/Cross-references: GB:U15155; NID:G603902; PIDN:AAA79912.1; PID:G603903
A/Experimental source: clone 1-P1
R/Mang, K.
submitted to the EMBL Data Library, September 1994
A/Reference number: S71155
A/Accession: S71155
A/Molecule type: mRNA
A/Residues: 1-102, 'A', '104-248 <MAN5>
A/Cross-references: EMBL:U15156; NID:G603904; PIDN:AAA79913.1; PID:G603905
A/Experimental source: clone 1-P38

C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-25/Domain: activation peptide #status predicted <APT>
F:26-246/Product: trypsin I #status predicted <MPT>
F:26-241/Domain: trypsin homology <TRY>
F:65,109,202/Active site: His, Asp, Ser #status predicted

Alignment Scores:

Score:	7.79e-26	Length:	248
Percent Similarity:	377.50	Matches:	88
Best Local Similarity:	53.06%	Conservative:	42
Query Match:	24.20%	Mismatches:	100
		Indels:	15
		Gaps:	4

US-10-037-270-482 (1-866) x S55067 (1-248)

```
Oy 42 ATGAAATAT-----GCTCTATTGGGTGCTCGTGGACATTTTCTTGT 92
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 1 MetLysPheLeuValLeuValAlaPheLeuGlyVal---AlaValAlaPheProIleSer 19

Oy 93 GACTCATCTGTTGAGAA-----GAGACCTCTCTCTAT 128
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 20 AspGluAspAspAspLysIleValGlyGlyTyrSerCysAlaArgSerAlaAlaProTyr 39

Oy 129 TTGGTGTACCTCAAGCTCATTCAACCCCTGTGGGCGTCCATCAAAACCAAGCTGG 188
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 40 GlnValSerLeuAsnSerGlyTyrHisPheCysGlyGlySerLeuIleSerSerGlnTyr 59

Oy 189 GTGGTGGCCCGGACCTCACTGCTATTACCAATGGAAGTGAAGTGAAGTGAAGTGAAG 248
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 60 ValLeuSerAlaAlaHisCysTyrLysSerSerIleGlnValLysLeuGlyGlnTyrAsn 79

Oy 249 AGCAGAGTCAGAGACGGTACTGAAACAGAAATTAACCCCATTCAGATCGTCCGCTACTGG 308
    ::|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 80 LeuAlaIaGlnAspGlySerGlnGlnThrIleSerSerSerValIleArgHisSer 99

Oy 309 AACATACATATAGGCCCCACAGATGACTCATGCTCATCAAGCTGCTAAGCTGCC 368
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 100 GlyTyrAsnSerAsnThrLeuAsnAsnAspIleMetLeuIleLysLeuSerLysAlaAla 119

Oy 369 ATGCTCATCCCAAAAGTCAGACCCCTCTCGCCACCAACATGTCAGGCCAGGCACT 428
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 120 ThrLeuAsnSerTyrValAsnThrValProLeuProThrSerCysValThrAlaGlyThr 139

Oy 429 GTCTGTACTCTCAGGTTTGGACTGAGCAAGAAAACAGTGGCCGACCTGACTTG 488
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 140 ThrCysLeuIleSerGlyTyrGlyAsnThrLeuSerSerGlySerLeuTyrProAspVal 159

Oy 489 CGGCAGAACCTGGAGGCCCGCTGATGCTGATGAGAAATGCCAAAAACAGAACAGGA 548
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 160 LeuGlnCysLeuAsnAlaProValLeuSerSerSerGlnCysSerSerAlaTyrProGly 179

Oy 549 AAAAGCCACAGAGATTCCTATGCTGAATTTTGGAAGTATTGACGCCGAATTTTGGG 608
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 180 ArgIleThrSerAsnMetIleCysIleGlyTyrLeuAsnGlyLysLysAspSerCysGln 199

Oy 609 GAGGTGGCCCTGTACTGTCTCATCTGCAAGAACAAGCTCCAGGGAATC-----GAG 659
    ::|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 200 GlyAspSerGlyLysProValValCysAsnGlyGlnLeuGlnGlyPheValSerTyrPly 219

Oy 660 GTGGGCACTTCATGGAGGGAGACCTCGGCATCTACCAATGTTTACCAATATATATCC 719
    ::|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 220 IleGlyCysAlaGlnLysGlyTyrProGlyValTyrThrLysValCysAsnTyrValSer 239

Oy 720 TGGATTGAGAACACT 734
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 240 TrpIleLysThrThr 244
```

RESULT 2
S33496
trypsin (BC 3.4.21.4) IV form a - human

C:Species: Homo sapiens (man)
C:Date: 03-Mar-1994 #sequence_revision 03-Aug-1995 #text_change 15-Aug-1997
C:Accession: S33496
R:Wiegand, U.; Corbach, S.; Mann, A.; Kang, J.; Mueller-Hill, B.
Submitted to the EMBL Data Library, March 1993

A:Description: Identification, cloning and characterization of a cDNA encoding a human B₂

A:Reference number: S33496

A:Accession: S33496

A:Molecule type: mRNA

A:Residues: 1-304 <MTE>

A:Cross-references: EMBL: X72781

C:Genetics:

A:Gene: GDB: PRSS4; TRY4

A:Cross-references: GDB: 335300

A:Map position: 7q35-7q35

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; serine proteinase

F:81-296/Domain: trypsin homology <TRY>

F:120,164,257/Active site: His, Asp, Ser #status predicted

Alignment Scores:

Pred. No.:	1.28e-23	Length:	304
Score:	353.00 <td>Matches:</td> <td>87</td>	Matches:	87
Percent Similarity:	51.38% <td>Conservative:</td> <td>43</td>	Conservative:	43
Best Local Similarity:	34.39% <td>Mismatches:</td> <td>113</td>	Mismatches:	113
Query Match:	22.63% <td>Indels:</td> <td>10</td>	Indels:	10
		Gaps:	4

US-10-037-270-482 (1-866) x S33496 (1-304)

```
Oy 6 TTATCCCTGAGCTAAGGGGAGAGCTGATGCATCATTAATATATGTC-----TTC 56
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 47 LeuHisProLeuLeuGlyGlyArgThrTyrArgAlaAlaArgAspAlaAspGlyCysGlu 66

Oy 57 TATTGGGTGTCTCGTGGGACATTTTCTTGCTGATCATCTGTTTCAG----- 107
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 67 AlaLeuGlyThrValAlaValAlaProPheAspAspAspLysIleValGlyGlyThr 86

Oy 108 ---AAAGAACCTGCTCCCTATTGCTGCTACTCAAGTCTCACTTCAACCCCTGTGTG 164
    ::|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 87 CysGlnGlnAsnSerLeuProTyrGlnValSerLeuAsnSerGlySerHisPheCysGly 106

Oy 165 GGCCTCTCATCAAAACCCAGCTGGGTGGTGGCCCGGACCTGCTACTGCTATTACCAATCTG 224
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 107 GlySerLeuIleSerGlnGlnThrValValSerAlaAlaHisCysTyrLysThrArgIle 126

Oy 225 AAAGTATGCTGGGAAATTTCAAGACAGAGTCAAGACGGTACTGAACAGCAATTAAC 284
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 127 GlnValArgLeuGlyGlyHisAsnIleValValLeuGlyGlnGlnPheIleAsn 146

Oy 285 CCCATTCAATGCTCCGCTACTGAGTACATGATGAGCCCGGACAGATGACTCATG 344
    ::|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 147 AlaAlaLysIleIleAsnHisProLysTyrAsnArgAspThrLeuAspAsnAspIleMet 166

Oy 345 CTCATCAAGCTGGGTAGCCGTGATGCTCATATCCAAAGTCCAGCCCTTCCCTGCCC 404
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 167 LeuIleLysLeuSerSerProAlaValIleAsnAlaArgValSerThrIleSerLeuPro 186

Oy 405 ACCACCAATGTCAGCCAGCAGTCTGTCTTACTCTCAGTGTGGAAGTGAAGCAAGAA 464
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 187 ThrLalProProAlaAlaGlyThrGlyCysLeuIleSerGlyTyrGlyAsnThrLeuSer 206

Oy 465 AACAGTGGCAGACCCCTGACTGGCAGAACTGGAAGCCCGGATGATCTGATCGA 524
    ::|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 207 PheGlyAlaAspTyrProAspGlnLeuLysCysLeuAspAlaProValLeuThrGlnAla 226

Oy 525 GAATGCCAAAAAAGCAAGAAAGCAAGAAAGCAAGAAATTCATATGATGAAATTTGTG 584
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 227 GluCysLysAlaSerTyrProGlyLysIleThrAsnSerMetPheCysValGlyPheLeu 246

Oy 585 AAAGTATTCAGCCGAATTTTGGGAGGTGGCCCTTCTACTGATCTGCAAGAACAG 644
    ::|||:::|||||:::|||||:::|||||:::|||||:::|||||:::
```

```

Db      247 GIUGLYGLYLSAPSERCYSGINARGASERGLYGLYPROVALVALCYSAENGLYGLN 266
Qy      645 CTCAGGAGCA--ATCGAGTGGGGCACTTCATGGGA-----GGGAGCGTCGGCATCTAC 695
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      267 LeuGlnIyValValSerTrpGlyHISGLYCYSAIATrPlySAEnArGProGlyValTyr 286
Qy      696 ACCAATGTTTACAAATATGATCTCGATTGAGAACACT 734
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      287 ThrLySValTyrAnTyrValAspTrpILeYSAEnpThr 299

RESULT 3
trypsin (EC 3.4.21.4) I precursor [validated] - human
N/Alternate names: trypsin, catenonic; trypsinogen I
C/Species: Homo sapiens (man)
C/Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 08-Dec-2000
C/Accession: A25852; B61066; A43988
R/Entl, M.; Nakamura, Y.; Ogawa, M.; Yamamoto, T.; Nishide, T.; Mori, T.; Matsubara, K.
Gene 41, 305-310, 1986
A/Title: Cloning, characterization and nucleotide sequences of two cDNAs encoding human
A/Reference number: A91544; MUID:86221712; PMID:3011602
A/Accession: A25852
A/Molecule type: mRNA
A/Residues: 1-247 <ENI>
A/Cross-references: GB:M22612; NID:9521215; PIDN:AAA61231.1; PID:9521216
R/Kimland, M.; Russick, C.; Marks, W.H.; Borgstrom, A.
Cln. Chim. Acta 184, 31-46, 1989
A/Title: Immunoreactive anionic and cationic trypsin in human serum.
A/Reference number: A61066; MUID:90091010; PMID:2598466
A/Accession: B61066
A/Molecule type: protein
A/Residues: 16-43 <KIM>
R/Koivunen, E.; Huhtala, M.L.; Stenman, U.H.
J. Biol. Chem. 264, 14095-14099, 1989
A/Title: Human ovarian tumor-associated trypsin. Its purification and characterization
A/Reference number: A43988; MUID:9340515; PMID:2503510
A/Accession: A43988
A/Molecule type: protein
A/Residues: 16-54 <KOI>
A/Experimental source: mucinous ovarian tumor cyst fluid
C/Genetics:
A/Gene: GDB:PRSS1; TRY1
A/Cross-references: GDB:119620; OMIM:276000
A/Map position: 7q35-7q35
A/Note: The human genome contains at least ten trypsin genes or pseudogenes, at least tw
C/Superfamily: trypsin; trypsin homology
C/Keywords: hydrolase; pancreas; phosphoprotein; protein digestion; serine proteinase; z
F/1-15/Domain: signal sequence #status predicted <SIG>
F/16-246/Product: trypsinogen I #status experimental <ZYM>
F/16-23/Domain: activation peptide #status predicted <ENZ>
F/24-246/Product: trypsin I #status predicted <APT>
F/24-239/Domain: trypsin homology <TRY>
F/30-160/48-64,139-206,171-185,196-220/Disulfide bonds: #status predicted
F/63,107,200/Active site: His, Asp, Ser #status predicted
F/75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Alignment Scores:
Pred. No.: 1.46-23 Length: 247
Score: 352.50 Matches: 82
Percent Similarity: 50.43% Conservative: 36
Best Local Similarity: 35.04% Mismatches: 109
Query Match: 22.60% Indels: 7
DB: Gaps: 2

US-10-037-270-482 (1-866) x A25852 (1-247)
Qy      54 TTCTATTGGGTGTCGTCGGAGACATTTTCTTGCTGACATGCTGTTAG----- 107
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      9 PheValAlaAlaAlaValAlaAlaProPheAspAspAspAspIleValAlGlyIlyTyr 28
Qy      108 -----AAAGAGACCTGCTCCCTATTGGTGTACTCAAGTCTCAAGTCTCAACCCCTGT 161
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      29 AsnCyseGIUGIuAEnSerValProTyrGlnValSerIeuAEnSerGlyTyrHISArpCyS 48

```

```

Qy      162 GTGGGCTCTCATCAACCCAGCTGGTGCTGGCCCCAGCTCATCTGTTATACCAAT 221
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      49 GlyIySerIeuIleEnGInGInTrpValValSerAlaGlyIySerIyrylYrlySerArg 68
Qy      222 CTGAAGATGCTGGGAAATTTCAAGACAGATCAGACAGCGTACTGAACAGCAATT 281
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      69 ILeGInValArgEnGlyGInHISAnIleGlnValLeuGInGlyAnGInGInPheIle 88
Qy      282 AACCCATTCAGATCGCTCCGCTACTGGAACACTACAGTCATGCCCCCAGCATGACCTC 341
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      89 AsnAlaAlaIySerIleAlaGHisProGInTyrAspArgIyThrIeuAAsnAspIle 108
Qy      342 ATGCTCATCAAGCTGGCTGAAGCTGCATGCTCAATCCCAAGTCCAGCCCTTCCCTC 401
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      109 MetIeuIleIySerIeuSerIerAlaValAlaIleAsnAlaArgValSerThrIleSerIeu 128
Qy      402 GCCACCAACCAATGTACAGCCAGCCACTGTCTGTCTACTCTCAGCTTGAAGACCAA 461
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      129 ProThrAlaArProAlaThrGlyThrIyScyLeuIleSerGlyTrpGlyAsnThrAla 148
Qy      462 GAAACAGTGGCCGACACCTGACTTGGCAGAACCTGAGGCCCTCCGATGTCTGAT 521
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      149 SerSerGlyAlaAspIyrrProAspGluLeuGInCySeuAspAlaProValLeuSerGln 168
Qy      522 CGAATGCGCAAAACAGAACAGAAAGCCACAGAACTTCCTTATGTGAAATTT 581
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      169 AlAyCyseGIuAlaSerTrpProGlyIySIlleThSerAsnMetPheCySeValGlyPhe 188
Qy      582 GTGAAGATATTCAGCCGAATTTTGGGAGGTGGCGGTGCTCATCTCATCTGCCAAGAC 641
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      189 LeuGInGlyIyLySAPSERCYSGInGlyIyAspSerGlyIyProValValCYSAENGLY 208
Qy      642 AACCTCAGGAGATCCAGGTC-----GGGACTTCATGGAGGGAGCTCGCATC 692
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      209 GlnLeuGlnIyValIyAlSerTrpGlyAspGlyCYSAIAGInIySAEnIyProGlyVal 228
Qy      693 TACACCAATGTTTACAAATATGATCTGATCTGATGAGAACACT 734
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      229 TyrThrLySValTyrAnTyrValIySerTrpILeYSAEnpThr 242

RESULT 4
B25528
trypsin (EC 3.4.21.4) precursor - mouse
C/Species: Mus musculus (house mouse)
C/Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 22-Jun-1999
C/Accession: B25528
R/Stevenson, B.J.; Hagenbuechle, O.; Wellauer, P.K.
Nucleic Acids Res. 14, 8307-8330, 1986
A/Title: Sequence organisation and transcriptional regulation of the mouse elastase II ar
A/Reference number: A93646; MUID:87066713; PMID:3641189
A/Accession: B25528
A/Molecule type: mRNA
A/Residues: 1-246 <STE>
A/Cross-references: GB:X04574; NID:954918; PIDN:CAA28243.1; PID:954919
C/Superfamily: trypsin; trypsin homology
C/Keywords: calcium binding; hydrolase; protein digestion; serine proteinase
F/1-23/Domain: signal sequence #status predicted <SIG>
F/24-246/Product: trypsin #status predicted <MRT>
F/30-160/48-64,132-233,139-206,171-185/Disulfide bonds: #status predicted
F/63,107,200/Active site: His, Asp, Ser #status predicted
F/75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Alignment Scores:
Pred. No.: 1.72e-23 Length: 246
Score: 351.50 Matches: 77
Percent Similarity: 53.49% Conservative: 38
Best Local Similarity: 35.81% Mismatches: 97
Query Match: 22.53% Indels: 3
DB: Gaps: 1

US-10-037-270-482 (1-866) x B25528 (1-246)

```


Db 201 ValValCysaenglyInLeugInglyValValserTrpGlyTrgIyCyAlaMeArg 220
 Oy 678 GGGAGCTGGCACTACCAATGTTTAAATATGATTCCTGGATTGGAACACT 734
 Db 221 AsnTrpProGlyValIyTrhIyValCysaenTrpAsnAlaTrpIleGlnAsnThr 239

RESULT 8

TRRT1

trypsin (EC 3.4.21.4) I precursor - rat

N:Alternate names: trypsinogen I

C:Species: Rattus norvegicus (Norway rat)

C:Date: 17-Dec-1992 #sequence_revision 17-Dec-1982 #text_change 24-Sep-1999

C:Accession: B22657; A00948

R:Craik, C.S.; Choo, O.L.; Swift, G.H.; Quinto, C.; MacDonald, R.J.; Rutter, W.J.

J. Biol. Chem. 259, 14255-14264, 1984

A:Title: Structure of two related rat pancreatic trypsin genes.

A:Reference number: A22657; PMID:85054880; PMID:6094547

A:Accession: B22657

A:Molecule type: DNA

A:Residues: 1-246 <CRA>

A:Cross-references: GB:J00778; NID:g206507; PIDN:AA98518.1; PID:g206508

A>Note: the authors translated the codon ATC for residue 6 as Leu and GAC for residue 17

R:MacDonald, R.J.; Stealy, S.J.; Swift, G.H.

J. Biol. Chem. 257, 9724-9732, 1982

A:Title: Two similar but nonallelic rat pancreatic trypsinogens. Nucleotide sequences of

A:Reference number: A00948; PMID:82265624; PMID:6896710

A:Accession: A00948

A:Molecule type: mRNA

A:Residues: 1-246 <MAC>

A:Cross-references: GB:J00778; NID:g206507; PIDN:AA98518.1; PID:g206508

C:Genetics:

A:introns: 14/1; 67/2; 152/1; 197/3

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolyase; pancreas; protein digestion; serine proteinase; zymogen

F:1-15/Domain: signal sequence #status predicted <SIG>

F:16-23/Domain: activation peptide #status predicted <APT>

F:24-246/Product: trypsin I #status predicted <ENZ>

F:24-239/Domain: trypsin homology <TRY>

F:30-160,48-64,132-233,139-206,171-185/Disulfide bonds: #status predicted

F:63,107,200/Active site: His, Asp, Ser #status predicted

F:75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Alignment Scores:

Pred. No.: 4,866-23 Length: 246

Score: 346.50 Matches: 76

Percent Similarity: 54.50% Conservative: 39

Best Local Similarity: 36.02% Mismatches: 93

Query Match: 22.21% Indels: 3

DB: 1 Gaps: 1

US-10-037-270-482 (1-866) x TRRT1 (1-246)

Oy 111 GAAGACCTGCTGCTATTTGGTGAACAGTCACTTCAACCCCTGTGGCGTC 170
 Db 32 GlnHservalProtyGlnValserLeuAsnSerGlyTrhAspHeCyGlyGlySer 51
 Oy 171 CTCATCAACCCAGTGGTGGTGGCCCCAGTCACTCTCTATTTCACAAATCGAAATG 230
 Db 52 LeuIleAsnAspGlnTrpValValserAlaIleAsnIleAsnIleValIleGlnVal 71
 Oy 221 ATGCTGGAAATTTCAAGACAGAGTCAAGACGCTGTAACACAAATTAACCCCAATT 290
 Db 72 ArgLeuGlyGlnIleAsnIleAsnValIleuGlyGlyAspGlyGlnPheIleAsnAla 91
 Oy 291 CAGATCTCCGCTGCTGAGTCACTACAGTCAAGAGCCGCCAAGATGACCTCATGCTCATC 350
 Db 92 LysIleIleLeuYnHsIleProAsnIleSerSerTrpTrhLeuAsnAsnAspIleIleLeuIle 111
 Oy 331 AAGCTGGCTAAAGCTGCTCATGCTCAATCCAAAGTCCAGCCCTTCCTCCGCCACACC 410
 Db 112 LysIleAsnSerSerProValIleLysLeuAsnAlaArgValAlaIleProValIleLeuProSerAla 131
 Oy 411 AATGTCAGGCGACGACTGTCTGCTACTCTAGCTTGAAGTGAAGCAAGAAACAGT 470

Db 132 CyAlaProAlaIyTrhGlnCysLeuIleSerGlyTrpIyAsnTrhLeuSerAsnGly 151
 Oy 471 GGGCGACCCCTGACTGGCGGCAACCTGGAGCCCGGTGAGTGCATCGAATGC 530
 Db 152 ValAsnAsnProAspLeuLeuGlnCysValAsnProValIleuSerGlnAlaAspCys 171
 Oy 531 CAAAGAACAGAAAGAAAGAAAGCCACAGAAATTCCTATGTGAAATTTGAAAGTA 590
 Db 172 GluAlaIleAlaTrpProGlyGlnIleTrhSerSerMetIleCysValGlyPheLeuGly 191
 Oy 591 TTGAGCGAATTTTGGGAGGTGGCCCTGCTACTGTCACTTGCAAGACCACTCAG 650
 Db 192 GlyIleAspSerCysGlnGlyAspSerGlyIleProValIyCysaenglyInLeuGln 211
 Oy 651 GGATC-----GAGGTGGGCACTTCATGGAGGAGCGATCGATCACCACAT 701
 Db 212 GlyIleValserTrpGlyTrgIyCyAlaLeuProAsnProGlyValIyTrhIy 231
 Oy 702 GTTACAAATATGATTCCTGATTCGATTCGAAACACT 734
 Db 232 ValCysaenPheValGlyTrpIleGlnAspThr 242

RESULT 9

TRDG

trypsin (EC 3.4.21.4) precursor, anionic - dog

N:Alternate names: cationic trypsinogen

C:Species: Canis lupus familiaris (dog)

C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 18-Jun-1999

C:Accession: A26273

R:Pinsky, S.D.; LaForge, K.S.; Scheele, G.

Mol. Cell. Biol. 5, 2669-2676, 1985

A:Title: Differential regulation of trypsinogen mRNA translation: full-length mRNA seque

A:Reference number: A26273; PMID:86284628; PMID:3841794

A:Accession: A26273

A:Molecule type: mRNA

A:Residues: 1-247 <PIN>

A:Cross-references: GB:M1589; NID:g164094; PIDN:AAA0899.1; PID:g164095

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolyase; pancreas; protein digestion; serine proteinase; zymogen

F:1-15/Domain: signal sequence #status predicted <SIG>

F:16-23/Domain: activation peptide #status predicted <APT>

F:24-247/Product: trypsin, anionic #status predicted <ENZ>

F:24-239/Domain: trypsin homology <TRY>

F:30-160,48-64,132-233,139-206,171-185/Disulfide bonds: #status predicted

F:63,107,200/Active site: His, Asp, Ser #status predicted

F:75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Alignment Scores:

Pred. No.: 5,396-23 Length: 247

Score: 346.00 Matches: 76

Percent Similarity: 52.56% Conservative: 47

Best Local Similarity: 32.48% Mismatches: 103

Query Match: 22.18% Indels: 8

DB: 1 Gaps: 2

US-10-037-270-482 (1-866) x TRDG (1-247)

Oy 57 TATTGGTGTCTGCTGGGACATTTTCTTCTGATCATGTGTTACG----- 107
 Db 9 PheLeuGlyAlaAlaValAlaIleProThrAspAspAspIleValGlyIyTr 28
 Oy 108 -----AAAGAGACCTGCTGCTATTTGGTGAACAGTCACTTCAACCCCGT 161
 Db 29 ThrCysGlnGluAsnSerValProtyGlnValserLeuAsnAlaGlyTrhAspHeCy 48
 Oy 162 GTGGAGCTGCTCATCAACCCAGCTGGTGGCCCCAGTCACTCTATTTCACAAAT 221
 Db 49 GlyIleSerLeuIleSerAspGlnTrpValValserAlaIleAsnIleCysGlyTrhSerArg 68
 Oy 222 CTGAAGTATGCTGGAAATTTCAAGACAGAGTCAAGACGCTGTAACACAAAT 281
 Db 69 IleGlnValArgLeuGlyGlyTrhAsnIleAspValIleuGlyGlyAsnGlyGlnPheIle 88

QY	282	AACCCCATCATGATCGTCCGTACTGGAATCACTAGACATGAGCCGCCACAGCATGACCTC	34.1
		:::::::::: :::::::::: :::::::::: :::::::::: ::::::::::	
Db	89	AamserAlaIyValIleIhrghIstProAbnityrAsnSerTrpIleuAspAsnIle	108.9
QY	342	ATGCTCATCAAGCTGGCTAAGCCTGCATGCTCAATCCCAAGTCGACCCCTTCCCTC	40.1
		:::::::::: :::::::::: :::::::::: :::::::::: ::::::::::	
Db	109	MettIuIleIyLeuSerSerProAlaIValIleuAsnAlaaryAlaIThrIleSerLeu	128.9
QY	402	GCCACCAACCAATGTGAGCCGAGCACTGTCGTCTACTCTGACAGTTTGACITGAGCCAA	46.1
		:::::::::: :::::::::: :::::::::: :::::::::: ::::::::::	
Db	129	ProAlaGAlaCyAlaIAlaIaProAlIyThrGInCySLeuIleSerIyTrpGlyAsnThrLeu	148.9
QY	462	GAAAACAGTGCAGCCGACCCCTGACTTGCGGACAAACCTGGAGGCCCCCGTATGTCTAT	52.1
		:::::::::: :::::::::: :::::::::: :::::::::: ::::::::::	
Db	149	SerSerGIyThrAsnTyxProGInLeuIleuGInCyLeuAspAlaProIleuThrGIn	168.9
QY	522	CGAGATGCCCAAAAACAAACAAGAAAAAGCCACAGCAATTCCTTATGTGAAATTT	56.1
		:::::::::: :::::::::: :::::::::: :::::::::: ::::::::::	
Db	169	AlaGInCySbIuAlaSerTyxProGInIyIleIhrGluAsnMetIleCySAlaGlyPhe	188.9
QY	582	GTGAAGAATTCACGCCGAATTTTTGGGAGAGTGCCCGTGTCTCATCTGCAAGAC	64.1
		:::::::::: :::::::::: :::::::::: :::::::::: ::::::::::	
Db	189	LeuGInGlyGlyLeuAspSerCySgInGlyAspSerGlyGlyProValAlaCySAsnGly	208.9
QY	642	AAGCTCCAGGGAATC-----GAGGTGGGCACTTCATGGAGAGGAGCTGGCATC	652.1
		:::::::::: :::::::::: :::::::::: :::::::::: ::::::::::	
Db	209	GluLeuGInGlyIleValSerTrpGlyTyrglyCySAlaGInIlyAsnIlySProGlyVal	228.9
QY	693	TACACCAATGTTTACAATATATATCTGCGATTTGAGAACACT	73.4
		:::::::::: :::::::::: :::::::::: :::::::::: ::::::::::	
Db	229	TyrThrIySValCySAsnIheValaSerTrpIleGInSerThr	242.9

RESULT 10
S12764
C:Synain (EC 3.4.21.4) III precursor - human
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 22-Jun-1999
C:Accession: S12764
R:Tani, T.; Kawashima, I.; Mita, K.; Takiguchi, Y.
Nucleic Acids Res. 18, 1631, 1990
A:Title: Nucleotide sequence of the human pancreatic trypsinogen III cDNA.
A:Reference number: S12764; MUID:90221895; PMID:2326201
A:Accession: S12764
A:Molecule type: mRNA
A:Residues: 1-247 <TAN>
A:Cross-references: EMBL:X15505; NID:g37459; P1DN:CAA33527.1; P1D:g37460
C:Genetics:
A:Gene: GDB:PRSS3; TRY3
A:Cross-references: GDB:335297
A:Map position: 7q35-7q35
C:Superfamily: trypsin; trypsin homology
C:Keywords: calcium binding; hydrolase; pancreas; protein digestion; serine proteinase.
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-21/Domain: activation peptide #status predicted <AP>
F:22-241/Product: trypsin III #status predicted <MAT>
F:24-239/Domain: trypsin homology <TRY>
F:30-160,48-64,139-206,171-185/Disulfide bonds: #status predicted
F:63,107,200/Active site: His, Asp, Ser #status predicted
F:75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted
Alignment Scores:

Pred. No.:	5,39e-23	length:	247
Score:	346.00	Matches:	79
Percent Similarity:	52.14%	Conservative:	43
Best Local Similarity:	33.76%	Mismatches:	104
Query Match:	22.18%	Indels:	8
DB:	2	Gaps:	3

US-10-037-270-482 (1-866) X SL2764 (1-247)
QY 57 TATTGGGNGTCCGTGACATTTTCTTGGCTGACTATCGTTCAG-----107
:::|||||

[illegible]

RESULT 11
5S5066
trypsin (EC 3.4.21.4) II precursor, pancreatic (clone 2-P29) - chicken
N/Alternate names: trypsinogen II
C/Species: Gallus gallus (chicken)
C/Date: 23-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 22-Jun-1999
C/Accession: 5S5066; S72347
R/Wang, K.; Gan, L.; Lee, I.; Hood, L.
Biochem. V. 307, 471-479, 1995
A>Title: Isolation and characterization of the chicken trypsinogen gene family.
A/Reference number: 5S5065; MUID:95251611; PMID:7731885
A/Accession: 5S5066
A/Molecule type: mRNA
A/Residues: 1-248 <MAN1>
A/Cross-references: EMBL:U01517; NID:g603906; PIDD:AA97914.1; PID:g603907
A/Experimental source: clone 2-P29
A/Accession: S72347
A/Molecule type: DNA
A/Residues: 1-248 <MAN2>
A/Cross-references: EMBL:U01517; NID:g603906; PIDD:AA97914.1; PID:g603907
A/Experimental source: clone 2-P29
A/Superfamily: trypsin; trypsin homology
C/Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F:1-16/Domain: signal sequence #stratus predicted <SIG>
F:17-25/Domain: activation peptide #stratus predicted <APT>
F:26-248/Product: trypsin II #stratus predicted <MAT>
F:26-241/Domain: trypsin homology <TR>

F:65,109,202/Active site: His, Asp, Ser #status predicted

Alignment Scores:

Pred. No.:	7,37e-23	Length:	248
Score:	344.50	Matches:	76
Percent Similarity:	54.03%	Conservative:	38
Best Local Similarity:	36.02%	Mismatches:	94
Query Match:	22.08%	Indels:	3
DB:	2	Gaps:	1

US-10-037-270-482 (1-866) x S55066 (1-248)

```

QY 111 GAGACCCCTGCTCCCTATTGGTGTACCTCAAGTCACTTCAACCCCTGTGGCGCTC 170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 34 GlnHisSerValProTyrGlnValSerLeuHisSerGlyTyrHisPheCysGlyGlySer 53
QY 171 CTGATCAAAACCCAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 54 LeuLeuHisSerGlnTyrValLeuSerValHisCysTyrTyrSerArgIleGlnVal 73
QY 231 ATGCTGGGAATTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 290
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74 ArgLeuGlyGlyTyrValHisLeuAspValGlnIleHisSerGlyValValArgSerSer 93
QY 291 CAGATCGTCCGCTAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 350
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 94 ValIleIleArgHisProTyrTyrSerSerIleThrLeuHisAsnAspIleMetLeuIle 113
QY 351 AAGCTGGCTAAGCTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 114 LysLeuHisSerAlaValIleGlyTyrSerIleAspIleGlnProIleAlaLeuProSerSer 133
QY 411 AATGTCAGCCGACCTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 470
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 134 CysAlaIysAlaGlyThrGlnCysLeuIleSerGlyTyrGlyAsnThrLeuSerHisGly 153
QY 471 GCGCGACACCTGACCTGCGCGACACCTGCGCGACACCTGCGCGACCTGATCGAATGCG 530
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 154 TyrAsnTyrProGlnLeuLeuGlnCysLeuHisAsnAlaProIleLeuSerAspGlnCys 173
QY 531 CAAAAACAGAACAGAAAAAGCAGCAAGAACTCTTATGTGTGAATTTGTGAAGTA 590
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 174 GlnIleValTyrProGlyAspIleThrSerAsnMetIleCysValGlyPheLeuGlnGly 193
QY 591 TTCACGCGAATTTTGGGAGAGTGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 650
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 194 GlyIysAspSerCysGlnGlyAspSerGlyProValValCysAsnGlyGlnLeuGln 213
QY 651 GGAATC-----GAGGTGGGCACTTCATGGAGGGGAGAGCTCGCGCATACCAAT 701
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 214 GlyIleValSerTyrPglYIleGlyCysAlaLeuHisSerGlyTyrProGlyValTyrThrIys 233
QY 702 GTTTACAAATATGTATCTGTGATTTGAGAACACT 734
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 234 ValCysAsnTyrValAspTyrIleGlnIleThr 244

```

RESULT 12

TRRT2 trypsin (EC 3.4.21.4) II precursor - rat

N:Alternate names: trypsinogen II
C:Species: Rattus norvegicus (Norway rat)

C:Date: 05-Apr-1993 #sequence_revision 30-Sep-1997 #text_change 18-Jul-1997

R:Accession: A22657; A00949

R:Cratk, C.S.; Choo, Q.L.; Swift, G.H.; Quinto, C.; MacDonald, R.J.; Rutter, W.J.

J. Biol. Chem. 259, 14255-14264, 1984

A:Title: Structure of two related rat pancreatic trypsin genes.

A:Reference number: A22657; MUID:85054880; PMID:6094547

A:Accession: A22657

A:Molecule type: DNA

A:Residues: 1-246 <CRA>

R:MacDonald, R.J.; Steary, S.J.; Swift, G.H.

J. Biol. Chem. 257, 9724-9733, 1982

A:Title: Two similar but nonallelic rat pancreatic trypsinogens. Nucleotide sequences of

A:Reference number: A00948; MUID:82265624; PMID:6896710

A:Accession: A00949

A:Molecule type: mRNA

A:Residues: 9-246 <MAC>

C:Comment: The trypsin II mRNA is present in much lower quantities than the trypsin I mRNA

C:Genetic:

A:Insertions: 14/1; 67/2

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen

F:1-15/Domain: signal sequence #status predicted <SIG>

F:16-23/Domain: activation peptide #status predicted <APR>

F:24-246/Product: trypsin II #status predicted <ENR>

F:24-239/Domain: trypsin homology <TRY>

F:30-160,48-64,132-233,139-206,171-185/Disulfide bonds: #status predicted

F:63,107,200/Active site: His, Asp, Ser #status predicted

F:75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Alignment Scores:

Pred. No.:	1.69e-22	Length:	246
Score:	340.50	Matches:	78
Percent Similarity:	50.83%	Conservative:	45
Best Local Similarity:	32.23%	Mismatches:	108
Query Match:	21.83%	Indels:	11
DB:	1	Gaps:	2

US-10-037-270-482 (1-866) x TRRT2 (1-246)

```

QY 42 ATGAAATATGCTCTTCTTTGGGTCGTCGAGACATTTTCTTGCTGACTCATCT 101
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 MetArgAlaLeuLeuPheLeuAlaLeuValGlyAlaAlaValAlaPheProValAspAsp 20
QY 102 GTTCAGAAA-----GAGACCCCTGCTCCCTATTGGTGTAC 137
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21 AspAspIysIleValGlyGlyTyrThrCysGlnGlnHisSerValProTyrGlnValSer 40
QY 138 CTCAAGTCTCACTTCAACCCCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 197
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 41 LeuHisSerGlyTyrHisPheCysGlyGlySerLeuHisAsnAspGlnTyrValValSer 60
QY 198 CCGCTCACTGCTATTACCAAAATCTGAAGTGTCTGGGAAATTTCAAGACAGAGTCT 257
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 AlaAlaHisCysTyrTyrSerArgIleGlnValArgLeuGlyGlnHisAsnIleAsnVal 80
QY 258 AGAACCGGTACTGAACAGACATTTAACCCATTGAGTGTCCGCTGATCGAATACAGT 317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 81 LeuGlnGlyAspGlnGlnPheIleAsnAlaAlaIysIleIleIysHisProAsnPheAsp 100
QY 318 CATAGCCGCCCAAGATGACCTCATGCTCATCAAGCTGGCTAAGCTGCCATCTCAAT 377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 101 ArgIysThrLeuHisAsnAspIleMetLeuIleIysLeuSerSerProValIysLeuAsn 120
QY 378 CCCAAATCCAGCCCTTCCCTCCGACCAACATGTCAGGCGCAGGCACTGTCTGCTA 437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 AlaArgValAlaThrValAlaLeuProSerSerCysAlaProAlaGlyThrGlnCysLeu 140
QY 438 CTCTCAGTTTGGATCGAGCCCAAGAAACAGTGGCCGACACCTTGCGGACAGAAC 497
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 141 IleSerGlyTyrPglYAsnThrLeuSerSerGlyValAsnGlnProAspLeuLeuGlnCys 160
QY 498 CTGAGGCCCCCGTGTATGTCTGATCGAATGCCAAATGCCAAATGCCAAATGCCAAATGCC 557
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 161 LeuAspAlaProLeuLeuProGlnAlaAspCysGlnIleAspTyrProGlyIysIleThr 180
QY 558 AGGAATTCCTTATGCTGGAATTTGTGAAGTATTCAGCGGAATTTTGGGAGAGTGGCC 617
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 AspAsnMetValCysValGlyPheLeuGlnGlyIysIysAspSerCysGlnGlyAspSer 200
QY 618 GTTCTACTGCTCATCTCAAAAGACACCTCCAGGGAATC-----GAGGTGGGCGAC 668
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 201 GlyIysProValValCysAsnGlyGlnLeuGlnIleValIleSerTyrPglYIysCys 220
QY 669 TTCATGGAGGAGGAGCTGGCATCTACCAATTTTAAATATGTATCTGATTTGAG 728
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db	221	AlbLeuProAspAsnProGlyValTyrThrIysValCysAsnTyrValAspTrpIleGln	240
QY	729	AAACACT	734
		:::	
Db	241	AspThr	242

Best Local Similarity: 35.38% Mismatches: 97
 Query Match: 21.44% Indels: 3
 DB: 1 Gaps: 1

US-10-037-270-482 (1-866) x B25852 (1-247)

```

QY 108 AAGAGACCTGCTGCTATTTGATGATCTCAAGTCTCAATCAACCCCTGTGGGC 167
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 31 GlutlunenservValProTyrGlnValSerleuanserGlyTyrHisPheCysGly 50
QY 168 GTCTCATCAAAACCCAGCTGGTGGTGGCCCACTGCTCATTTTACCAATCTGAA 227
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 51 SerleuileserGlnTyrValValSerValGlnHisCysTyrTyrSerArgIleGln 70
QY 228 GTGATGCTGGGAATTTCAAGACAGAGCTGAGAGCTGATCAAGCAATTAACCCC 287
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 71 ValArgLeuGlyGlnHisValenIleGlnValLeuGlnGlnGlnPheIleHisAla 90
QY 288 ATTCAGATCTGCTGCTGCTGATCACTACAGTACAGCCCAAGATGATGCTGCTC 347
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 91 AlalysleileleArgHisProTyrTyrSerArgTyrHisLeuAspAsnAlaIleleu 110
QY 348 ATCAAGCTGCTAAGCTGCTGCTGCTCAATCCCAAGTCCAGCCCTTCCCTGCCACC 407
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 111 IleTyrLeuSerSerProAlaValIleAsnSerArgValSerAlaIleSerLeuProthr 130
QY 408 ACCAATGTCAGCCAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 467
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 131 AlalProAlaAlaGlyThrGlnSerleuileserGlyTyrGlyAsnThrLeuSerSer 150
QY 468 AGTGGCCGACACCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 527
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 151 GlnValAspTyrProAspGlnLeuGlnCysLeuAspAlaProValLeuSerGlnAlaGln 170
QY 528 TGCCAAAAACAGAACAGAAAGAAAGCCACAGAAATTCCTATGTGTGAAATTTGTGAA 587
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 171 CysGlnAlaSerTyrProGlyTyrIleThrAsnAsnMetPheCysValGlyPheLeuGln 190
QY 588 GTATTCACCCGAATTTTGGGGAGGTGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 647
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 191 GlnGlyTyrAspSerCysGlnGlyAspSerGlyTyrProValValSerAsnGlyGlnLeu 210
QY 648 CAGGAGATC-----GAGGTGGGCGCATTCATGAGAGGAGCGTGGCATCTACACC 698
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 211 GlnGlyIleValSerTyrGlyTyrCysAlaGlnGlyAsnArgProGlyValTyrThr 230
QY 699 AATGTTTCAAAATATGATCTGCTGATGAGAACT 734
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 231 LysValTyrAsnTyrValAspTyrIleTyrAspThr 242

```

RESULT 15

S13813
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 02-Dec-1993 #sequence_revision 03-Aug-1995 #text_change 22-Jun-1999
 C:Accession: S13813
 A:Reference number: S13813; MUID:91065383; PMID:1701147
 A:Accession: S13813
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-247 <HDB>
 A:Cross-references: EMBL:X54703, NID:9829, PIDN:CAA38513.1, PID:9830
 C:Superfamily: trypsin; trypsin; trypsin homology
 C:Keywords: hydrolase; protein digestion; serine proteinase
 F:24-239/Domain: trypsin homology <TRY>
 F:63,107,200/Active site: His, Asp, Ser #status predicted

Alignment Scores: 1.66e-21 Length: 247
 Pred. No.: 329.50 Matches: 77
 Score:

Percent Similarity: 53.33% Conservative: 43
 Best Local Similarity: 34.22% Mismatches: 96
 Query Match: 21.12% Indels: 9
 DB: 2 Gaps: 2

US-10-037-270-482 (1-866) x S13813 (1-247)

```

QY 69 CTGCTGGAGCAATTTTCTTTGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 128
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 24 IleValGlyGlyTyrThrCysAlaGlnValSerVal-----ProTyr 37
QY 129 TTGCTGATCTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 188
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 38 GlnValSerLeuAsnAlaGlyTyrHisPheCysGlyGlySerLeuIleAsnAspGlnTyr 57
QY 189 GTGCTGGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 248
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 58 ValValSerAlaAlaHisCysTyrGlnTyrHisIleGlnValArgLeuGlyGlyTyrAsn 77
QY 249 AGCAGACTGAGAGCGTACTGAAACAGACAAATTAACCCATTCAGATGCTGCTGCTG 308
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 78 IleAspValLeuGlnGlyGlnGlnPheIleAspAlaSerTyrIleArgHisPro 97
QY 309 AACTACATTAAGCGCCCAAGATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 368
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 98 LysTyrSerSerTyrThrLeuAspAsnAspIleLeuLeuIleTyrLeuSerTyrProAla 117
QY 369 ATGCTCAATCCCAAGTCCAGCCCTTCCCTGCGCCCAAGTCCAGCCAGGCACT 428
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 118 ValIleAsnAlaArgValSerThrLeuLeuLeuProSerAlaCysAlaSerAlaGlyThr 137
QY 429 GTCTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 488
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 138 GlnCysLeuIleSerGlyTyrGlyAsnThrLeuSerSerGlyValAsnTyrProAspLeu 157
QY 489 CGGCAAGACTGAGGCGCCCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 548
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 158 LeuGlnCysLeuValAlaProLeuLeuSerHisAlaAspCysGlnAlaSerTyrProGly 177
QY 549 AAAAGCCAGGAATTCCTATGCTGAAATTTGTGAAAGATTCAGCCGAATTTTGGG 608
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 178 GlnIleThrAsnAsnMetIleCysAlaGlyPheLeuGlnGlyTyrLysAspSerCysGln 197
QY 609 GAGGTGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 198 GlnAspSerGlyTyrProValAlaCysAsnGlyGlnLeuGlnGlyIleValSerTyrGly 217
QY 660 GTGGGCGCATTCATGAGAGGAGCGTGGCATCTACACCAATGTTTCAATATGATCC 719
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 218 TyrGlyCysAlaGlnGlyLysProGlyValTyrThrLysValCysAsnTyrValAsp 237
QY 720 TGGAATGAGAACT 734
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 238 TrpIleGlnGlnThr 242

```

Search completed: November 8, 2003, 02:12:47
 Job time : 36 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 8, 2003, 00:18:41 ; Search time 17.5 Seconds
(without alignments)
4654,300 Million cell updates/sec

Title: US-10-037-270-482

Perfect score: 1560
Sequence: 1 ggcactactccctgagctca.....tgggaatgtagcatcactagc 866

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 255726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+n2p,model -DEV=xjh
-Q=/cgn2_1/USPRO.spool/US10037270/runat_07112003_140514_8746/app.query.fasta.1.1031
-DB=SwissProt 41 -QFMT=fastan -SUFFIX=isp -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcrc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10037270.QCGN_1_1_1@runat_07112003_140514_8746 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-NO TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	ID	Description
1	376.5	24.1	248	TRY2_CHICK	Q90628 gallus gall
2	375.5	24.1	248	TRY1_CHICK	Q90627 gallus gall
3	353	22.6	304	TRY3_HUMAN	P35030 homo sapien
4	352.5	22.6	247	TRY1_HUMAN	P07477 homo sapien
5	351.5	22.5	246	TRY2_MOUSE	P07146 mus musculu
6	350.5	22.5	246	TRY1_CANPA	P06871 canis famul
7	349	22.4	243	TRY1_XENLA	P19799 xenopus lae
8	346.5	22.2	246	TRY1_XENLA	P00762 rattus norv
9	346	22.2	247	TRY1_RAT	P06872 canis famul
10	344.5	22.1	248	TRY3_CHICK	Q90629 gallus gall
11	340.5	21.8	246	TRY2_RAT	P00763 rattus norv
12	335.5	21.5	231	TRY2_PIG	P00768 sus scrofa
13	334.5	21.4	247	TRY2_HUMAN	P07478 homo sapien
14	331.5	21.2	244	TRY2_XENLA	P70059 xenopus lae
15	329.5	21.1	247	TRY2_BOVIN	P08426 bos taurus
16	328.5	21.1	247	TRY3_RAT	P08426 bos taurus
17	312.5	20.0	243	TRY1_BOVIN	P00760 bos taurus
18	312	20.0	229	TRY1_SQUAC	P00764 squatus aca

19	312	20.0	247	TRY4_RAT	P12788 rattus norv
20	311.5	20.0	238	TRY3_SALSA	P35033 salmo salar
21	311	19.9	246	TRYB_RAT	P32832 rattus norv
22	309	19.8	246	TRYA_RAT	P32831 rattus norv
23	287.5	18.4	244	TRY6_HUMAN	Q92876 homo sapien
24	285	18.3	231	TRY2_SALSA	P35032 salmo salar
25	274.5	17.9	242	TRY1_SALSA	P35031 salmo salar
26	274.5	17.6	256	TRY1_HUMAN	Q92875 homo sapien
27	269.5	17.3	241	TRY1_GADMO	P16049 gadus morhu
28	260	16.7	241	TRYX_GADMO	Q91041 gadus morhu
29	253	16.2	251	TRY6_HUMAN	Q92876 homo sapien
30	252	16.2	260	TRY3_HUMAN	Q92875 homo sapien
31	247.5	15.9	255	CATG_HUMAN	P08311 homo sapien
32	244	15.6	248	GRLL_RAT	Q06605 rattus norv
33	243	15.6	260	GRLL_MOUSE	Q06605 rattus norv
34	237	15.2	277	NRPN_MOUSE	Q06605 rattus norv
35	236.5	15.2	277	NRPN_HUMAN	Q06605 rattus norv
36	235.5	15.1	261	CATG_MOUSE	Q28233 mus musculu
37	235	15.1	248	GRAC_MOUSE	P28233 mus musculu
38	232	14.9	250	TRY1_HUMAN	Q92875 homo sapien
39	231	14.8	248	NRPI_RAT	P18291 rattus norv
40	228.5	14.6	247	GRAB_MOUSE	P04187 mus musculu
41	221.5	14.2	254	TRY4_HUMAN	Q92875 homo sapien
42	221	14.2	261	TRY3_MOUSE	P07526 mus musculu
43	220	14.1	248	TRY3_HUMAN	Q92875 homo sapien
44	219	14.0	246	TRY1_RAT	P97592 rattus norv
45	218	14.0	258	VSPI_AGRAC	Q91831 agkistrodon

ALIGNMENTS

RESULT 1

TRY2_CHICK ID TRY2_CHICK STANDARD; PRT; 248 AA.

AC Q90628; DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Trypsin I-P38 precursor (EC 3.4.21.4).

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archaeoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Pancreas;

RX MEDLINE=95251611; PubMed=7733885;

RA Wang K., Gan L., Lee I., Hood L.E.;

RT "Isolation and characterization of the chicken trypsinogen gene

family."

RL Biochem. J. 307:471-479 (1995).

CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg|-Xaa, Lys|-Xaa.

CC -1- SUBCELLULAR LOCATION: Extracellular.

CC -1- TISSUE SPECIFICITY: HIGH LEVELS ARE SEEN IN THE PANCREAS WHILE

LOWER LEVELS ARE FOUND IN THE LIVER, SPLEEN AND THYMUS.

-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for connected
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sb-sib.ch).

EMBL: U15156; AAA79913.1; -
PIR: S55067; S55067.
HSSP: P00763; IDPO.
MEROPS: S01.258; -
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00089; trypsin_1.
DR SMART: SM00020; Tryp_Spc; 1.

Alignment Scores:

Pred. No.: 2,056-26 Length: 248
 Score: 375.50 Matches: 87
 Percent Similarity: 53.06% Conservative: 43
 Best Local Similarity: 35.51% Mismatches: 100
 Query Match: 24.07% Indels: 15
 Gaps: 4

US-10-037-270-482 (1-866) x TRY1_CHICK (1-248)

```

QY 42 ATGAAATAT-----GTCTTATTTGGGCTCTCGTGGACATTTTCTTGTCT 92
Db 1 MetLysPheLeuValLeuValAlaPheValGlyValThrVal--AlaPheProIleSer 19
QY 93 GACATCACTGTTCGAAA-----GAAAGCCCTGCTCCCTAT 128
Db 20 AspGluAbpAapLysIleValGlyGlyTyrSerCysAlaArgSerAlaAlaProTyr 39
QY 129 TTGGTGTACCTCAAGTCTCACTTCAACCCCTGTGGGCGCTCATCAAAACCCAGCTGG 188
Db 40 GluValSerLeuAenSerGlyTyrHisPheCysGlySerLeuIleSerGlnTyr 59
QY 189 GTGCTGCCCCAGCTCACTGCTATTACCAAAATGAAAGTGAAGTGGGAAATTTCAG 248
Db 60 ValLeuSerAlaAlaHisCysTyrLysSerSerIleGlnValLysLeuGlyGluTyrAsn 79
QY 249 AGCAGATCAGAGACGCTACTGAGACAGACATTAACCCATTCAGATCGCTGCTACTGG 308
Db 80 LeuAlaAlaGlnAbpGlySerGlnThrIleSerSerLysValIleArgHisSer 99
QY 309 AACTACATCATACGCGCCCGCCAGACATGACCTCAAGTCAAGCTGAGCTGAGCTGCC 368
Db 100 GlyTyrAsnAlaAsnThrLeuAsnAsnAspIleMetLeuIleLysLeuSerLysAlaAla 119
QY 369 ATGCTCAATCCCAAGTCCAGCCCTTCCCTCCGACCAACCAATGTCAGCCAGCACT 428
Db 120 ThrLeuAenSerTyrValAlaAsnThrValProLeuProThrSerCysValThrAlaGlyThr 139
QY 429 GTCTGTACTCTTACGTTTGGACTGAGCCCAAGAAAACATGAGCCGACCTGACTTG 488
Db 140 ThrCysLeuIleSerGlyTyrGlyAsnThrLeuSerSerGlySerLysTyrProAspVal 159
QY 489 CGGCAAGACCTGAGGCGCCCGCTGATGCTGATGAGAAATGCCAAAAAGAACAGACAGA 548
Db 160 LeuGlnCysLeuAsnAlaProValLeuSerSerGlnCysSerSerAlaTyrProGly 179
QY 549 AAAAGCCACAGAAATTCCTTATGTGTGAATTTGTGAAGATTCAGCCGAATTTTGGG 608
Db 180 ArgIleThrSerAsnMetIleCysGlyIleGlyTyrLeuAsnGlyLysAbpSerCysGln 199
QY 609 GAGGTGCGCTGCTACTGCTGATCTGCAGAACAGAGCTCCAGGGAATC-----GAG 659
Db 200 GlyAspSerGlyGlyProValValCysAsnGlyGlnLeuGlnGlyIleValSerTyrGly 219
QY 660 GTGGGAGCACTTCATGAGGAGGAGCGTGGGATTCACCAATGTTTCAATATGATGCC 719
Db 220 IleGlyCysAlaGlnLysGlyTyrProGlyValTyrThrLysValCysAenTyrValSer 239
QY 720 TGGATTGAGAACT 734
Db 240 TripleLysThrThr 244
  
```

RESULT 3

TRY3_HUMAN STANDARD; PRT; 304 AA.
 AC P35030; P15951; Q15665; O9UOV; 1
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Trypsin III precursor (EC 3.4.21.4) (Brain trypsinogen)
 DE (Mecrotrypsinogen) (Trypsin IV).
 GN PRSS3 OR TRY3 OR PRSS4 OR TRY4.
 OS Homo sapiens (Human).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS A AND B).
RC TISSUE=Brain;
RX MEDLINE=9412394; PubMed=8294000;
RA Wiegand U., Corbach S., Mann A., Kang J., Mueller-Hill B.;
RT "Cloning of the cDNA encoding human brain trypsinogen and
RL characterization of its product."
RL Gene 136:167-175(1993).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM C).
RC TISSUE=Pancreas;
RX MEDLINE=90221895; PubMed=2326201;
RA Tani T., Kawashima I., Mita K., Takiguchi Y.;
RT "Nucleotide sequence of the human pancreatic trypsinogen III cDNA."
RL Nucleic Acids Res. 18:1631-1631(1990).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM C).
RA Fukuda S.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) (ISOFORM A).
RX MEDLINE=21686310; PubMed=11827488;
RA Katona G., Berglund G.I., Hajdu J., Graf L., Szilagyi L.;
RT "Crystal structure reveals basis for the inhibitor resistance of human
RL brain trypsin."
RL J. Mol. Biol. 315:1209-1218(2002).
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-Xaa, Lys-Xaa.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=A;
CC IsoId=P35030-1; Sequence=Displayed;
CC Name=B;
CC IsoId=P35030-2; Sequence=VSP_005409;
CC Name=C;
CC IsoId=P35030-3; Sequence=VSP_005410;
CC -1- TISSUE SPECIFICITY: Pancreas and brain.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC
CC -----
CC EMBL: X72781; CAB58178.1; -
CC EMBL: X71345; CA50484.1; -
CC EMBL: X15505; CA13527.1; -
CC EMBL: D45417; BAA08257.1; -
CC PIR: S33496; S33496.
CC PDB: 1H4W; 15-FEB-02.
CC MEROPS: S01.174; -.
CC GeneW: HGNC:9486; PRSS3.
CC GO: GO:0005576; C:extracellular; NAS.
CC GO: GO:0004295; F:trypsin activity; NAS.
CC GO: GO:0006508; P:proteolysis and peptidolysis; NAS.
CC InterPro: IPR001314; Chymotrypsin.
CC InterPro: IPR001254; Ser_protease_Try.
CC Pfam: PF00089; Trypsin_1.
CC PRINTS: PR00722; CHYMOTRYPSIN.
CC SMART: SM00020; TRYPSIN_DOM_1.
CC PROSITE: PS00240; TRYPSIN_DOM_1.
CC PROSITE: PS00134; TRYPSIN_HIS_1.
CC PROSITE: PS00135; TRYPSIN_SER; FALSE_NEG.
CC HydroLase: Serine protease; Digestion; Zymogen; Signal;
CC MultiGene family; Alternative splicing; 3D-structure.
CC SIGNAL 1 ?
CC PROPEP ? 80 ACTIVATION PEPTIDE.
  
```


Oy	282	AACCCATTGAGATGTGCCTCGCATGGAACTCAAGTCATAGAGGCCCAAGAGTGAATC	341
Db	89	AsnAlaAlaIysIleIleArgHisProGlnTyPheArgLysThrLeuAsnAsnSerPile	108
Oy	342	ATGTCATCAAGCTGGCTAAGCCTGCATGCTCAATCCCAAAGTCAGCCCCTTC	401
Db	109	MetLeuIleIysLeuSerSerArgValIleAsnAlaArgAlaSerThrlleSerLeu	128
Oy	402	GCCACCACCAATGTGAGCGCACGACTGTCTGTACTCAAGGTTTGAGACTGGACCCA	461
Db	129	ProThrAlaProProAlaThrIleIynTrIysCySLeuIleSerLeIYrpgIysnHraIa	148
Oy	462	GAACAACAGTGGCCGACACCTTGACTTGGCGCAAACTGAGAAGCCCCGTGATGTGAT	521
Db	149	SerSerGIyAlaAspTyPheArgpGluLeuGlnCySLeuAspAlaProValIleuSerGln	168
Oy	522	CGAGAAATGCCAAAAACAGAACAGAAAAGCACAGAAATCTTATGTGTGAATAATTT	581
Db	169	AlaIysCySGluAlaSerTyPheGlyIysIleThrSerAsnMetPheCySValIGlyPhe	188
Oy	582	GTGAAGAATTAACGCCGAATTTTTGGGGAGAGTGCCCGTGTCTACTGTCACTGCAAGAC	641
Db	189	LeuGIuGIyGIyLyAspSerCySGlnGIyAspSerGIyGIyProValValCySAsnGIy	208
Oy	642	AAGCTCCAGGAATCAGAGTG-----GGGCACTTCATGGAGGAGGAGCTGGGCATC	692
Db	209	GlnLeuGIuGIyValValSerTrpGIyAspGIyCySAlaGIuIysAsnIysProGIyVal	228
Oy	693	TACACCAATGTTTACAAATATGATTCCTGGATTTGTAAGAAACT	734
Db	229	TyrTrnIysValIyAsnTyPValIyStrIleIySAsnThr	242

ID	TRY2_MOUSE	STANDARD:	PRT:	246 AA.
AC	P07146;			
DT	01-APR-1998 (Rel. 07, Created)			
DT	01-APR-1998 (Rel. 07, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Trypsin II, anionic precursor (EC 3.4.21.4) (Pretrypsinogen II).			
GN	TRIZ.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_Taxid=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=A/J;			
RX	MEDLINE=87066713; PubMed=3641189;			
RA	Stevenson B.J., Hagenbuchle O., Wellauer P.K.;			
RT	"Sequence organisation and transcriptional regulation of the mouse			
RL	elastase II and trypsin genes.";			
RL	Nucleic Acids Res. 14:8307-8330(1986).			
CC	-1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-Xaa, Lys-Xaa.			
CC	-1- SUBCELLULAR LOCATION: Extracellular.			
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X04574; CAA28243.1; -.			
DR	EMBL; X04577; CAA28245.1; -.			
DR	PIR; B25528; B25528.			
DR	HSSP; P00763; IDPO.			
DR	MEROPS; S01.064; -.			
DR	MGI; MGI:102759; Try2.			
DR	InterPro; IPR001314; Chymotrypsin.			

DR	Interpro: IPRO01254; Ser protease_Try.
DR	Pfam: PF000689; trypsin; 1.
DR	PRINTS; PRO0722; CHYMOTRYPSIN.
DR	SMART; SMO0020; Tryp_Spc; 1.
DR	PROSITE; PSS00240; TRYPsin_DOM; 1.
DR	PROSITE; PS00134; TRYpsin_HIS; 1.
DR	PROSITE; PS00135; TRYpsin_SER; 1.
KW	Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
FT	SIGNAL 1 15
FT	PROPEP 16 22 ACTIVATION PEPTIDE.
FT	CHAIN 24 246 TRYPSIN II, ANIONIC.
FT	ACT_SITE 63 63 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 107 107 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 200 200 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	DISULFID 30 160 BY SIMILARITY.
FT	DISULFID 48 164 BY SIMILARITY.
FT	DISULFID 132 233 BY SIMILARITY.
FT	DISULFID 139 206 BY SIMILARITY.
FT	DISULFID 171 185 BY SIMILARITY.
FT	DISULFID 196 220 BY SIMILARITY.
FT	SITE 194 194 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
SO	SEQUENCE 246 AA; 26203 MW; CERB6C9TAAAC2D07AD CRC64;

Alignment Scores:	
Pred. No.:	3,19e-24
Score:	351.50
Percent Similarity:	53.49%
Best local Similarity:	35.81%
Query Match:	22.53%
DB:	1
US-10-037-270-482 (1-866) x TRY2_MOUSE (1-246)	
Length:	246
Matches:	77
Conservative:	38
Mismatches:	97
Indels:	3
Gaps:	1

[illegible]

QY 699 ATGTTTACAAATATGATCTGATTGAGAACCTGCTAAGAC 743
 Db 231 LyvalCysaenTyValaSPTrIleGlnaenThrIleAlaSP 245

RESULT 6

TRY1_CANFA STANDARD; PRT; 246 AA.

DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Trypsin, cationic precursor (EC 3.4.21.4).
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 NC NCB1_TaxID=9615;

RT [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=86284628; PubMed=3841794;
 RX Pinsky S.D., Laforge K.S., Scheele G.;
 RA "Differential regulation of trypsinogen mRNA translation: full-length
 RT mRNA sequences encoding two oppositely charged trypsinogen isoenzymes
 in the dog pancreas";
 RL Mol. Cell. Biol. 5:2669-2676(1985).
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg|-Xaa, Lys|-Xaa.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; M11590; AAA30900.1; -.
 DR PIR; B26273; TRDGC.
 DR HSP; P00761; 1EPT.
 DR MEROPS; S01.151.-.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam; PF00089; trypsin_1.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PSS0240; TRYPsin_DOM; 1.
 DR PROSITE; PS00134; TRYPsin_HIS; 1.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 KW Hydrolyase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
 KW Multigene family.
 FT SIGNAL 1 15
 FT PROPEP 16 23 ACTIVATION PEPTIDE.
 FT CHAIN 24 246 TRYPSIN, CATIONIC.
 FT ACT_SITE 63 63 CHANGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 107 107 CHANGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 200 200 CHANGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 30 160 BY SIMILARITY.
 FT DISULFID 48 64 BY SIMILARITY.
 FT DISULFID 132 233 BY SIMILARITY.
 FT DISULFID 139 206 BY SIMILARITY.
 FT DISULFID 171 185 BY SIMILARITY.
 FT DISULFID 196 220 BY SIMILARITY.
 FT SITE 194 194 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
 SQ SEQUENCE 246 AA; 26170 MW; E9E5A1DE2391BBB CRC64;

Alignment Scores:

Pred. No.: 3.94e-24 Length: 246
 Score: 350.50 Matches: 81
 Percent Similarity: 50.83% Conservative: 42
 Best Local Similarity: 33.47% Mismatches: 108
 Query Match: 22.47% Indels: 11
 DB: 1 Gaps: 2

US-10-037-270-482 (1-866) x TRY1_CANFA (1-246)

QY 42 ATGAATATGATCTTATTTGGGTGCTGCTGGACATTTTCTTTGCTGCTCATCT 101
 Db 1 MetIythrPheIlePheLeuAlaLeuGlyAlaThrValAlaPheProIleAsp 20
 QY 102 GTTCAGAA-----GAAACCTTCTCTTCTTTGGTCTAC 137
 Db 21 AspAspIleValGlyGlyTrpCysSerArgAsnSerValProIleValSer 40
 QY 138 CTCAGTCTCATTCACCCCTGTGGGGCTGCTCATCAACCCAGCTGGTCTGGCC 197
 Db 41 LeuAsnSerGlyTrpHisPheCysGlyGlySerLeuIleAsnSerGlnTrpValSer 60
 QY 198 CCAGCTCATCTTATTTTACCAATCTGAAAGTATGATGCTGGAAATTTCAAGACAGTC 257
 Db 61 AlaAlaIleCysGlyTrpSerArgIleGlnValArgLeuGlyGlyTrpAlaIleVal 80
 QY 258 AGAGAGGATCTGAACAGACATTTAACCCCATTCAGATGCTCGCTACTGAACTACAGT 317
 Db 81 SerGluGlyGlyGluGlnPheIleAsnAlaIleValSerIleArgHisProArgTrpAsn 100
 QY 318 CAAAGGCCCCACAGATGACCTTCATGCTCATGAGCTGCTAAGCTGCTGCTCAAT 377
 Db 101 AlaAsnThrIleAspAsnAspIleMetLeuIleValSerSerProAlaThrLeuAsn 120
 QY 378 CCCAAGTCCAGCCCTTCCCTGCGCCACCAATGTCAGGCGACGCTGCTGCTA 437
 Db 121 SerArgValSerAlaIleAlaLeuProIleValSerCysProAlaIleGlyThrGlnCysLeu 140
 QY 438 CTCTCAGTTTGGACTGAGCCCAAGAAACAGTGGCCGACCTTGCCTGCGCAGAAC 497
 Db 141 IleSerGlyTrpIleAsnThrGlnSerIleGlyIleAsnTrpProAspValLeuGlnCys 160
 QY 498 CTGAGGCCCCCTGATGCTGATTCAGATGCGCAAAAACAGAACAGAAAAAGCCAC 557
 Db 161 LeuIleAlaProIleLeuSerAspSerValCysArgAsnAlaTrpProIleGlnIleSer 180
 QY 558 AGGAATCTCTATGCTGGAATTTGGAATGATTCAGCCGAATTTTGGGAGGTGGCC 617
 Db 181 SerAsnMetCysLeuGlyTrpMetGluGlyValAspSerCysGlnGlyAspSer 200
 QY 618 GTTGCTACTGCTCATTCGCAAGCAAGCTTCAGGGAATC-----GAGTGGGGCAC 668
 Db 201 GlyIleProValValCysAsnGlyGluLeuGlnGlyValValSerTrpIleAlaGlyCys 220
 QY 669 TTATGAGGAGGAGCAGCTGCGCATCTACCAATGTTTCAAAATATGATCTGATGAG 728
 Db 221 AlaGlnIleGlyLeuSerProIleValSerProIleValCysIleTrpValSerTrpIleGln 240
 QY 729 AACACT 734
 Db 241 GlnThr 242
 RESULT 7
 TRY1_XENLA STANDARD; PRT; 243 AA.
 ID TRY1_XENLA
 AC P19759;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Trypsin precursor (EC 3.4.21.4).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodinae; Xenopus.
 NC NCB1_TaxID=8355;
 RN SEQUENCE FROM N.A.
 RP TISSUE=Pancreas;
 RX MEDLINE=91007255; PubMed=2210372;
 RA Shi Y.B., Brown D.D.;
 RT "Developmental and thyroid hormone-dependent regulation of pancreatic
 RT genes in Xenopus laevis";
 RT Genes Dev. 4:1107-1113(1990).

CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-Xaa, Lys-Xaa.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ch/announce/>
 CC or send an email to license@ebi.ebi.ac.uk).

DR EMBL: X53458; CA37538.1; -
 DR PIR: A35871; A35871.
 DR HSSP: P00763; IDPO.
 DR MEROPS: S01.151; -
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam: PF00089; trypsin_1.
 DR PRINTS: PRO0722; CHYMOTRYPSIN.
 DR SMART: SM00090; TRYP_SPC; 1.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KM Hydroxylase; Serine protease; Digestion; Zymogen; Signal;
 KM Multigene family.
 FT SIGNAL 1 15 BY SIMILARITY.
 FT PROPEP 16 20 ACTIVATION PEPTIDE.
 FT CHAIN 21 243 TRYPSIN.
 FT ACT_SITE 60 60 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 104 104 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 197 197 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 27 157 BY SIMILARITY.
 FT DISULFID 45 61 BY SIMILARITY.
 FT DISULFID 129 230 BY SIMILARITY.
 FT DISULFID 136 203 BY SIMILARITY.
 FT DISULFID 168 182 BY SIMILARITY.
 FT DISULFID 193 217 BY SIMILARITY.
 FT SITE 191 191 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
 SQ SEQUENCE 243 AA; 25492 MW; C5B8345A8B3F8031 CRC64;

Alignment Scores:
 Pred. No.: 5.39e-24 Length: 243
 Score: 349.00 Matches: 78
 Percent Similarity: 51.88% Conservative: 46
 Best Local Similarity: 32.64% Mismatches: 107
 Query Match: 22.37% Indels: 8
 DB: 1 Gaps: 2

US-10-037-270-482 (1-866) x TRY1_XENLA (1-243)

QY 42 ATGAATAATGCTCTTCTTTGGGCTGCTGCGGACATTTTCTTGGCTACTCATCT 101
 DB 1 MetlyshphenleuLeuCyValleuLeuClValAlaAlaAlaPheAspAspLys 20
 QY 102 GTT-----CAGAAAGAACCCCTGCTCCCTATTGGTGGTACTCAAGTCT 146
 DB 21 IieIleelgylAlaIhrCyAlaIylserserValProTyrIleValserIleuAsnSer 40
 QY 147 CACTTCAACCCCTGTGGGCTGCTCATCAACCCAGCTGGTGTGCCCCAGCTCAC 206
 DB 41 GlyIyrHisPheCyGylGylSerIleuIleIhrAsnGlnIhrValIserAlaIleHis 60
 QY 207 TGCATTTTACCAATCTGAAGATGATGCTGGGAAATTTCAAGACAGATAGAGAGGT 266
 DB 61 CysTyrIylsAlaSerIleGlnValIArgIleuGylGylIleIleAsnIleAlaLeuSerGluGly 80
 QY 267 ACTGAACAGCAATTAACTCCCATTCAGATGCTGCGCTACGGAATCAAGTCAATAGCGCC 326
 DB 81 ThrIleGlnPheIleSerSerSerIylValIleArgIleSerGylTyrAsnSerTyrThr 100
 QY 327 CCAAGAGATGACCTCATGCTCATCAAGCTGAGCTGAGCTGCATGCTCAATCCCAAGTC 386
 DB 327 CCAAGAGATGACCTCATGCTCATCAAGCTGAGCTGAGCTGCATGCTCAATCCCAAGTC 386

DB 101 LeuAspAsnAspIleMetIleuIleLysLeuSerSerProAlaSerIleuAsnAlaVal 120
 QY 387 CAGCCCTTCCCTCGCGCACACCAATGTCAGCCAGGACACTGCTGTACTCTCAGGT 446
 DB 121 AsnThrValProLeuProSerGlyCySerAlaIleGlyIhrSerIleuIleSerGly 140
 QY 447 TTGAGCTGAGCAAGAAACACAGTGGCCGACACCTGACTTGGCGAGAACCTGAGGCC 506
 DB 141 TrpGlyIAsnThrIleuSerAsnGlySerAsnTyrProAspIleuGlnCysIleAsnAla 160
 QY 507 CCCGTGATGCTCATGAGAAATCCCAAAAAACGAAACAGAAAGAAACCCACAGAAATCC 566
 DB 161 ProIleuThrAsnAlaGlnCysAsnSerAlaTyrProGlyIleIleIhrAlaAsnMet 180
 QY 567 TTATGTGTGAATTTGGAAGATTCAGCCGATTTTGGGAGGTGGCGTGTCTACT 626
 DB 181 IleCysValIleIlyrIhrCglGylGylLysAspSerCysGlnIleLysAspSerGlyPro 200
 QY 627 GTCATCTGCAAGACAAACAGCTCCAGGAAATC-----GAGGTGGCGCATTCATGGA 677
 DB 201 ValValCysAsnGlyIleuGlnGlyValIserTyrGlyIleCysAlaMetArg 220
 QY 678 GGGAGCTGCGCATCTACACCAATGTTTACAAATATGATCTCGATTGAGAACT 734
 DB 221 AsnTyrProGlyValTyrThrIylsValCysAsnTyrAsnAlaIhrIleGlnAsnThr 239

RESULT 8
 TRY1_RAT STANDARD; PRT; 246 AA.
 ID TRY1_RAT
 AC P00762;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Trypsin I, anionic precursor (EC 3.4.21.4) (Pretypsinogen I).
 GN TRY1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RA MEDLINE=82265624; PubMed=6896710;
 RT McDonald R.J., Stary S.J., Swift G.H.;
 RT "Two similar but nonallelic rat pancreatic trypsinogens. Nucleotide
 RT sequences of the cloned cDNAs".
 RL J. Biol. Chem. 257:9724-9732(1982).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=85054880; PubMed=6094547;
 RA Craik C.S., Choo Q.L., Swift G.H., Quinto C., McDonald R.J.,
 RA Rutter W.J.;
 RT "Structure of two related rat pancreatic trypsin genes".
 RL J. Biol. Chem. 259:14255-14264(1984).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RC MEDLINE=87292123; PubMed=8112942;
 RA Sprang S., Standing T., Pieterick R.J., Stroud R.M., Finer-Moore J.,
 RA Xiong N.-H., Hamlin R., Rutter W.J., Craik C.S.;
 RT "The three-dimensional structure of Asn102 mutant of trypsin: role of
 RT Asp102 in serine protease catalysis".
 RL Science 237:905-909(1987).
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-Xaa, Lys-Xaa.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- MISCELLANEOUS: THIS SEQUENCE REPRESENTS THE PRECURSOR OF THE MAJOR
 CC FORM OF TRYPSIN PRODUCED BY THE ADULT PANCREAS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

RT mRNA sequences encoding two oppositely charged trypsinogen isoenzymes
 RT in the dog pancreas.";
 RL Mol. Cell. Biol. 5:2669-2676(1985).
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M11589; AAA30899.1; -.
 DR PIR: A26273; TRDG.
 DR HSSP: P00763; IDPO.
 DR MEROPS: S01.258; -.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam: PF00089; trypsin; 1.
 DR SMART: SM00020; Tryp_SPC; 1.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KW Hydroxylase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
 KW Multigene family.
 FT SIGNAL 1 15
 FT PROPEP 16 23
 FT CHAIN 24 247
 FT ACT_SITE 63 63
 FT ACT_SITE 107 107
 FT ACT_SITE 200 200
 FT DISULFID 30 160
 FT DISULFID 48 64
 FT DISULFID 132 233
 FT DISULFID 139 206
 FT DISULFID 171 185
 FT DISULFID 196 220
 FT SITE 194 194
 SQ SEQUENCE 247 AA; 26423 MW; 374B9D31D6DB8EAF CXC64;
 Alignment Scores:
 Pred. No.: 1.02e-23 Length: 247
 Score: 346.00 Matches: 76
 Percent Similarity: 52.56% Conservative: 47
 Best Local Similarity: 32.48% Mismatches: 103
 Query Match: 22.18% Indels: 8
 Gaps: 2
 US-10-037-270-482 (1-866) x TRY2_CANFA (1-247)
 QY 57 TATTGGGTGCTCCGCGGAGCATTTTCTTGTGCTGATCTGTTTCAG----- 107
 Db 9 PheLeuGlyAlaAlaValAlaThrProThraSparSparSylLeValGlyTyr 28
 QY 108 -----AAAGAGAGCCCTGCTCCCTATTGGTGATGCAAGTCACTTAAACCCCTG 161
 Db 29 ThrCysGluGluAsnSerValProTyrGlnValSerLeuAsnAlaGlyTyrHisPheCys 48
 QY 162 GTGGGCGCTCTCATCAACCCAGCTGGTGGTGGCCCGAGTCACTGTTATTCAAAAT 221
 Db 49 GlyGlySerLeuIleSerAspGlnTyrAlaValSerAlaAlaHisCysTyrLysSerArg 68
 QY 222 CTGAAAGTGAATGCTGGGAAATTTCAAGACAGAGTCAAGACCGTACTGAAACAGCAAT 281
 Db 69 IleGlnValArgLeuGlyGlnTyrAsnIleAspValLeuGluGlnGlyAsnGlnPheIle 88
 QY 282 AACCCATTGAGTGCCTGCGCTGAGCAACTAGCAAGTCAAGTCAAGCCGCCACAGAGTGAAT 341
 Db 89 AsnSerAlaLysValIleArgHisIleProGlnTyrAsnSerTyrPheLeuAspAsnSylLe 108
 QY 342 ATGCTCATCAAGCTGGCTGAAGCTGCGCTGCTCATGCTCAATCCCAAGTCCAGCCCTTCCCTC 401

Db 109 MetLeuIleLysLeuSerProAlaValLeuAsnAlaArgValAlaThrIleSerLeu 128
 QY 402 GCCACACCAATGTCAGGACGAGCTGTGCTACTCTCAGTGTGACAGCCAA 461
 Db 129 ProArgAlaCysAlaAlaPheGlyThrGlnCysLeuIleSerLysTyrPheAsnThrLeu 148
 QY 462 AAAACAGTGGCGACACCTGACTTGGCGAGACCTGAGGCCCCCGTATGTCAT 521
 Db 149 SerSerTyrThrAsnTyrProGlnLeuGlnCysLeuAspAlaProIleLeuThrGln 168
 QY 522 CGAATGTCACAAAACAGAACGAAAGAACGACGAAATTCCTATGTCGAATTT 581
 Db 169 ALGlnCysGlnAlaSerTyrProGlyGlnIleThrGlnAsnMetIleCysAlaGlyPhe 188
 QY 582 GTGAAAGTATTCAGCCGAAATTTTGGGAGGTGGCCGCTGCTACTGTCATCTGCAAAAGAC 641
 Db 189 LeuGlnIleGlyLysAspSerCysGlnGlyAspSerCysLysProValValCysAsnGly 208
 QY 642 AAGCTCCAGGAATC-----GAGGTGGGCACTTCATGAGAGGAGCGTGGCATC 692
 Db 209 GluLeuGlnGlyLeuValSerTyrGlyTyrGlyCysAlaGlnLysAsnLysProGlyVal 228
 QY 693 TACACCATGTTTACAAATATGTATCTGATGAGTGAACACT 734
 Db 229 TyrThrLysValCysAsnPheValAspTyrPheIleGlnSerThr 242
 RESULT 10
 TRY3_CHICK
 ID TRY3_CHICK STANDARD; PRT; 248 AA.
 AC Q90629;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Trypsin II-P29 precursor (EC 3.4.21.4).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=95251611; PubMed=773385;
 RA Wang K., Gan L., Lee I., Hood L.E.;
 RT "Isolation and characterization of the chicken trypsinogen gene
 family." J. 307:471-479(1995).
 RL Biochem. J. 307:471-479(1995).
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: HIGH LEVELS ARE SEEN IN THE PANCREAS WHILE
 CC LOWER LEVELS ARE FOUND IN THE LIVER, SPLEEN AND THYMUS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U15157; AAA79914.1; -.
 DR PIR: S55066; S55066.
 DR HSSP: P00763; IDPO.
 DR MEROPS: S01.151; -.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam: PF00089; trypsin; 1.
 DR SMART: SM00020; Tryp_SPC; 1.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KW Hydroxylase; Serine protease; Digestion; Pancreas; Zymogen; Signal;

DR PDB; 1F5R; 04-JUL-01.
DR PDB; 1F7Z; 04-JUL-01.
DR PDB; 1FV8; 04-JUL-01.
DR PDB; 1J14; 11-FEB-03.
DR PDB; 1J15; 11-FEB-03.
DR PDB; 1J16; 11-FEB-03.
DR PDB; 1J17; 11-FEB-03.
DR PDB; 1QL9; 25-AUG-00.
DR PDB; 1TRM; 15-JUL-93.
DR PDB; 2TRM; 16-JUL-88.
DR PDB; 3TCK; 04-JUL-01.
DR MEROS; S01.258; -.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR SMART; SMO0020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydroxylase; Serine protease; Digestion; Pancreas; Zymogen; signal;
KW Multigene family; 3D-structure.
FT SIGNAL 1 15
FT PROPEP 16 23 ACTIVATION PEPTIDE.
FT CHAIN 24 246 TRYPSIN II, ANIONIC.
FT ACT_SITE 63 63 CHARGE RELAY SYSTEM.
FT ACT_SITE 107 107 CHARGE RELAY SYSTEM.
FT ACT_SITE 200 200 CHARGE RELAY SYSTEM.
FT DISULFID 30 160
FT DISULFID 48 64
FT DISULFID 132 233
FT DISULFID 139 206
FT DISULFID 171 185
FT DISULFID 196 220
FT SITE 194 194
FT CONFLICT 84 84
FT CONFLICT 88 88
FT STRAND 25 25
FT STRAND 28 29
FT TURN 32 33
FT TURN 36 37
FT TURN 38 42
FT STRAND 46 54
FT TURN 55 56
FT STRAND 57 60
FT HELIX 62 64
FT STRAND 70 73
FT STRAND 77 77
FT TURN 78 79
FT STRAND 86 95
FT TURN 97 98
FT TURN 101 103
FT TURN 105 106
FT STRAND 109 113
FT STRAND 120 120
FT TURN 121 122
FT STRAND 123 123
FT STRAND 127 127
FT TURN 135 136
FT TURN 138 143
FT STRAND 157 157
FT STRAND 159 164
FT HELIX 168 174
FT TURN 176 178
FT TURN 181 182
FT STRAND 183 185
FT STRAND 194 194
FT TURN 197 198
FT TURN 200 201
FT STRAND 203 206
FT TURN 207 208
FT STRAND 209 216
FT TURN 223 224
FT STRAND 228 231
FT HELIX 232 234

REQUIRED FOR SPECIFICITY (BY SIMILARITY).
N -> D (IN REF. 1).
V -> I (IN REF. 1).

FT HELIX 236 246
SQ SEQUENCE 246 AA; 26228 MW; A8D3630809AE606 CRC64;
Alignment Scores:
Pred. No.: 3.23e-23
Score: 340.50
Percent Similarity: 50.83%
Best Local Similarity: 31.82%
Query Match: 21.83%
DB: 1
Gaps: 2
US-10-037-270-482 (1-866) x TRY2_RAT (1-246)
QY 42 ATGAATATCTCTTCAATTTGGGTGCTGCGGACATTTTCTTGCTCATCT 101
Db 1 MetArgAlaLeuLeuPheLeuAlaLeuValGlyAlaAlaValAlaPheProValAsp 20
QY 102 GTTCAGAA-----GAGACCTCTCTCTTATTTGGTGTAC 137
Db 21 AspAspLysIleValGlyGlyTyrThrCysGlnGlnLysSerValProTyrGlnValSer 40
QY 138 CTGAGCTCACTTCAACCCCTGTGGCGCTCTCATCAACCCAGCTGGCTGGCC 197
Db 41 LeuAsnSerGlyTyrHisPheCysGlyGlySerLeuIleAsnAspGlnTyrValSer 60
QY 198 CCACTCACTGATTTTACCAATCTGAAGATGCTGGGAATTTCAAGACAGATC 257
Db 61 AlaAlaHisCysTyrTyrSerArgIleGlnValArgLeuGlyGlnHisAsnIleAsnVal 80
QY 258 AGAGACGGTACTGAACAGACAAATTAAACCCATTCAAGTCTCGCTACTGAAC 317
Db 81 LeuGlnGlyAsnGlnGlnPheValAsnAlaAlaIleIleValSerHisPheAsp 100
QY 318 CATAGCCCCCAGAGATGACCTCATGCTCATGAAGTGGCTTAAGCTCGCATCTCAAT 377
Db 101 ArgLysThrLeuAsnAsnAspIleMetLeuIleLysLeuSerSerProValLysLeuAsn 120
QY 378 CCCAAGTCCAGCCCTTCCCTGCGCACCAATGTCAAGCCAGGCACTGTCTGCTA 437
Db 121 AlaArgValAlaThrValAlaLeuProSerSerCysAlaProAlaGlyThrGlnCysLeu 140
QY 438 CTCTCAGTTTGGACTGGACCAAGAAACAGTGGCCGACACCTTGACTGGCGAGAAC 497
Db 141 IleSerGlyTyrGlyAsnThrLeuSerSerGlyValAsnGlnProAspLeuLeuGlnCys 160
QY 498 CTGAGGCCCCCGTGAATGTGTGATCGAATGCCAAATAACAGAAACAGAAAGAGCCAC 557
Db 161 LeuAspAlaProLeuLeuProGlnAlaAspCysGlnLysSerTyrProGlyLysIleThr 180
QY 558 AGGAATTCCTTATGTGGAATTTGTGAAGTATTCAGCCGAATTTTGGGAGGTGGCC 617
Db 181 AspAsnMetValCysValGlyPheLeuGlnGlyGlyLysAspSerCysGlnGlyAspSer 200
QY 618 GTTGCTACTGTCACTGCAAGACAGACTCCAGGGAATC-----GAGTGGGGCAC 668
Db 201 GlyGlyProValValCysAsnGlyLeuGlnGlyIleValSerTyrGlyTyrCys 220
QY 669 TTGATGGAGGAGGACGTCGCATCTACCAATGTTTCAAAATGATCTTGATTTGAG 728
Db 221 AlaLeuProAspAsnProGlyValTyrThrLysValCysAsnTyrValAspTyrIleGln 240
QY 729 AACACT 734
Db 241 AspThr 242
RESULT 12
ID TRYP_PIG STANDARD; PRT; 231 AA.
AC P00761;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Trypsin precursor (EC 3.4.21.4).

OS Sus scrofa (pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN
 RP SEQUENCE OF 1-10.
 RA Charles M., Rovey M., Guidoni A.A., Desnuelle P.;
 RT "on trypsinogen and trypsin of pig.";
 RL Biochim. Biophys. Acta 69:115-122(1963).
 RN
 RP SEQUENCE OF 9-231.
 RA MEDLINE=73258692; Pubmed=4738933;
 RX Herndon M.A., Ericsson L.H., Neurath H., Walsh K.A.;
 RT "determination of the amino acid sequence of porcine trypsin by
 RT sequenator analysis.";
 RL Biochemistry 12:3146-3153(1973).
 RN
 RN X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
 RX MEDLINE=93187998; Pubmed=8445634;
 RA Huang Q., Liu S., Tang Y.;
 RT "Refined 1.6-A resolution crystal structure of the complex formed
 RT between porcine beta-trypsin and MCTI-A, a trypsin inhibitor of the
 RT equaen family. Detailed comparison with bovine beta-trypsin and its
 RT complex.";
 RL J. Mol. Biol. 229:1022-1030(1993).
 RN
 RN X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
 RX MEDLINE=92201369; Pubmed=1551419;
 RA Huang Q., Liu S., Tang Y., Zeng F., Qian R.;
 RT "Amino acid sequencing of a trypsin inhibitor by refined 1.6 A X-ray
 RT crystal structure of its complex with porcine beta-trypsin.";
 RL FEBS Lett. 297:143-146(1992).
 RN
 RN X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
 RX MEDLINE=95035057; Pubmed=7947985;
 RA Huang Q., Wang Z., Li Y., Liu S., Tang Y.;
 RT "Refined 1.8-A resolution crystal structure of the porcine epsilon-
 RT trypsin.";
 RL Biochim. Biophys. Acta 1209:77-82(1994).
 RN
 RN X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF COMPLEX WITH LDTI.
 RX MEDLINE=97390427; Pubmed=9242660;
 RA Stubbs M.T., Morenweiser R., Sturzebecher J., Bauer M., Bode W.,
 RA Huber R., Piechotka G.P., Matschner G., Sommerhoff C.P., Fritz H.,
 RA Auerwald E.A.;
 RT "The three-dimensional structure of recombinant leech-derived
 RT trypsin inhibitor in complex with trypsin. Implications for the
 RT structure of human mast cell trypsin and its inhibition.";
 RL J. Biol. Chem. 272:19931-19937(1997).
 RN
 RN X-RAY CRYSTALLOGRAPHY (2.03 ANGSTROMS) OF COMPLEX WITH LDTI.
 RX MEDLINE=98046095; Pubmed=9384562;
 RA di Marco S., Priestle J.P.;
 RT "Structure of the complex of leech-derived trypsin inhibitor (LDTI)
 RT with trypsin and modeling of the LDTI-trypsin system.";
 RL Structure 5:1465-1474(1997).
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-Xaa, Lys-Xaa.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 DR PDB; 1MCT; 31-JAN-94.
 DR PDB; 1AKS; 12-FEB-97.
 DR PDB; 1EPT; 07-FEB-95.
 DR PDB; 1TFX; 21-JAN-98.
 DR PDB; 1LDT; 20-MAY-98.
 DR PDB; 1ANI; 01-JUL-98.
 DR PDB; 1AVW; 18-NOV-98.
 DR PDB; 1C9P; 26-SEP-01.
 DR PDB; 1D3O; 30-NOV-99.
 DR PDB; 1DF2; 26-SEP-01.
 DR PDB; 1EJA; 02-MAR-01.
 DR PDB; 1EWJ; 28-JUN-00.

DR PDB; 1FMG; 07-NOV-01.
 DR PDB; 1FNE; 07-NOV-01.
 DR PDB; 1FNT; 07-NOV-01.
 DR PDB; 1LT2; 05-JUN-02.
 DR PDB; 1QOU; 14-JUN-00.
 DR MEROPS; S01.151; -.
 DR InterPro; IPR001254; Ser.protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS50240; TRYPsin_DOM; 1.
 DR PROSITE; PS00134; TRYPsin_HIS; 1.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 KM Hydroxylase; Serine protease; Digestion; Pancreas; Zymogen;
 KW 3D-structure.
 FT PROPEP 1 8
 FT CHAIN 9 231
 FT ACT_SITE 48 48
 FT ACT_SITE 92 92
 FT ACT_SITE 185 185
 FT DISULFID 15 145
 FT DISULFID 33 48
 FT DISULFID 117 218
 FT DISULFID 124 191
 FT DISULFID 156 170
 FT DISULFID 181 205
 FT SITE 179 179
 FT VARIANT 20 20
 FT STRAND 10 10
 FT STRAND 13 14
 FT TURN 17 18
 FT TURN 21 22
 FT STRAND 23 27
 FT STRAND 31 39
 FT TURN 40 41
 FT STRAND 42 45
 FT STRAND 47 49
 FT HELIX 55 58
 FT STRAND 62 62
 FT STRAND 63 64
 FT STRAND 71 80
 FT TURN 82 83
 FT TURN 86 88
 FT TURN 90 91
 FT STRAND 94 98
 FT STRAND 112 112
 FT TURN 120 121
 FT STRAND 123 128
 FT STRAND 142 142
 FT STRAND 144 150
 FT STRAND 153 159
 FT HELIX 161 163
 FT TURN 166 167
 FT STRAND 168 171
 FT TURN 174 175
 FT STRAND 179 179
 FT TURN 182 183
 FT TURN 185 186
 FT STRAND 188 191
 FT TURN 192 193
 FT STRAND 194 202
 FT STRAND 207 207
 FT TURN 208 209
 FT STRAND 210 210
 FT STRAND 212 216
 FT HELIX 217 219
 FT HELIX 221 231
 SQ SEQUENCE 231 AA; 24409 MW; A0A125CF7FC138C2 CRC64;
 Alignment Scores:
 Pred. No.: 9.14e-23
 Score: 335.50
 Percent Similarity: 52.89%
 Best Local Similarity: 34.67%
 Length: 231
 Matches: 78
 Conservative: 41
 Mismatches: 97
 ACTIVATION PEPTIDE.
 TRYPsin.
 CHARGE RELAY SYSTEM.
 CHARGE RELAY SYSTEM.
 CHARGE RELAY SYSTEM.
 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
 I -> V.

Query Match: 21.51% Indels: 9
DB: 1 Gaps: 2

US-10-037-270-482 (1-866) x TRY2_PIG (1-231)

QY 69 CTCCTGGGACATTTTCTTGTGCTGACTCATGTTCAAGAAAGAACCCCTGCTCCTAT 128
DB 9 TLeValGlyGlyTyrThrCysAlaAlaAsnSerIle-----ProTyr 22
QY 129 TTGGTGAACCTCAAGTCTCACTTCAACCCCTGGGGGCTTCATCAACCCAGCTGG 188
DB 23 GlnValSerLeuAsnSerGlySerHisPheCysGlyGlySerLeuIleAsnSerGlnTyr 42
QY 189 GTGCTGGCCCCGCTCACTGCTATTTCACCAATGGAAGTGAAGTCTGGGAATTTCAAG 248
DB 43 ValValSerAlaAlaHisCysTyrLysSerArgIleGlnValAlaGlyGlnHisAsn 62
QY 249 AGCAGAGTCAGAGACGGTACTGAAACAGACATTAACCCCATTCAGATGTCGCTACTGG 308
DB 63 TLeAspValLeuGlnGlyAsnGlnGlnPheIleAsnAlaAlaLysIleIleThrHisPro 82
QY 309 AACATCACTCATACCGCCCAAGAGATGACTCATGCTCATACCGCTGAACCTGACC 368
DB 83 AsnPheAsnGlyAsnThrLeuAspAsnAspIleMetLeuIleLysLeuSerSerProAla 102
QY 369 ATGCTCAATCCCAAGTCCAGCCCTTCCCTCCGCAACCAATGTCAGGCGCAGCACT 428
DB 103 ThrLeuAsnSerArgValAlaThrValSerLeuProAsnSerCysAlaAlaGlyThr 122
QY 429 GTCTGTCTACTCTCAAGTTTGGACTGAGCCAGAAAGAAACAGTGGCCGACCTGACTTG 488
DB 123 GlnCysLeuIleSerGlyTyrGlyAsnThrLysSerSerGlySerSerTyrProSerLeu 142
QY 489 CGGAGAACCTGGAGGAGCCCGCTGATGCTGATGAGAAATGCCAAAAGAACAGAACAGA 548
DB 143 LeuGlnCysLeuLysAlaProValLeuSerAspSerSerCysLysSerSerTyrProGly 162
QY 549 AAAAGCCACAGAAATTCCTTATGTGTGAATTTGTGAAGATTCACCGCAATTTTGGG 608
DB 163 GlnIleThrGlyAsnMetIleCysValAlaLysPheLeuGlnGlyLysAspSerCysGln 182
QY 609 GAGGTGGCCGTGTCTACTGTCATCTGCAGAACAGACGTCACGGAATC-----GAG 659
DB 183 GlyAspSerGlyGlyProValValCysAsnGlnGlnLeuGlnGlyIleValSerTyrPro 202
QY 660 GTGGGGCACTTCAGGGAGGGAGCGTCCGATCTACCAATGTTTACAAATATGATCC 719
DB 203 TyrGlyCysAlaGlnLysAsnLysProGlyValTyrThrLysValCysAsnTyrValAsn 222
QY 720 TGGATTGGAACACT 734
DB 223 TrpIleGlnGlnThr 227

RESULT 13
TRY2_HUMAN STANDARD: PRT; 247 AA.

AC P07478;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Trypsin II precursor (BC 3.4.21.4) (Anionic trypsinogen).
GN PRSS2 OR TRY2 OR TRY2P.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID:9606;
OX [1]
RN RP
SEQUENCE FROM N.A.
RX MEDLINE=86221712; PubMed=3011602;
RA Emi M., Nakamura Y., Ogawa M., Yamamoto T., Nishide T., Mori T.,
RA Matsubara K.,
RT "Cloning, characterization and nucleotide sequences of two cDNAs
RT encoding human pancreatic trypsins.",

RL Gene 41:305-310(1986).
RN [2]
RP SEQUENCE OF 16-49.
RX MEDLINE=90091010; PubMed=2598466;
RA Kimland M., Ruesick C., Marks W.H., Borgstrom A.,
RT "Immunoreactive anionic and cationic trypsin in human serum";
RL Clin. Chim. Acta 184:31-46(1989).
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg|-Xaa, Lys|-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ch/announce/>
CC or send an email to license@ebi.ebi.ac.uk).

DR EMBL/M27602; AAA61232.1; -.
DR PIR; B25852; B25852.
DR HSP; P00763; IDPO.
DR MEROPS; S01.258; -.
DR GeneW; HGNC:9483; PRSS2.
DR MIM; 601564; -.
DR GO; GO:0004295; F:trypsin activity; TAS.
DR InterPro; IPR001254; Ser protease_Try.
DR Pfam; PF00069; trypsin_1.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
KW Multigene family.
FT SIGNAL 1 15
FT PROPEP 16 23
FT CHAIN 24 247
FT ACT_SITE 63 63
FT ACT_SITE 107 107
FT ACT_SITE 200 200
FT DISULFID 30 160
FT DISULFID 48 64
FT DISULFID 171 185
FT DISULFID 196 220
FT SITE 194 194
FT SITE 194 194
SQ SEQUENCE 247 AA; 26488 MW; 82B0F41EB8E3D5B CRC64;

Alignment Scores:
Pred. No.: 1.14e-22 Length: 247
Score: 334.50 Matches: 75
Percent Similarity: 52.83% Conservative: 37
Best Local Similarity: 35.38% Mismatches: 97
Query Match: 21.44% Indels: 3
DB: 1 Gaps: 1

US-10-037-270-482 (1-866) x TRY2_HUMAN (1-247)

QY 108 AAAAAGAACCTGCTCCTTGTGCTGACTCATGCTCAACCCCTGCTGCGG 167
DB 31 GlnIleThrGlyAsnMetIleCysValAlaLysPheLeuGlnGlyLysAspSerCysGln 50
QY 168 GTCTCATCAACCCAGCTGGTGGTGGCCCCGCTCACTGCTATTTCACCAATTTGAAA 227
DB 51 SerLeuIleSerGlnGlnThrProValValSerLeuAsnSerGlyTyrHisPheCysGly 70
QY 228 GTGATGCTGGGAATTTTCAGAGCAGAGTCAAGACGCTGTAACAGACATTAACCCC 287
DB 71 ValArgLeuGlyGlnHisAsnIleGlnValLeuGlnGlnGlnPheIleAsnAla 90
QY 288 ATTGAGTGTGCTGCTGCTGAGTCAATGATGAGCCGCCACAGATGACCTCATCTC 347
DB 91 AlaLysIleIleArgHisProLysTyrAsnSerArgThrLeuAspAsnAspIleLeu 110

QY	348	ATCAAGCTGGGCTACAGCCCTGGCATGTGCTCAATCCCAAGTCCAGCCCTCCCTGGCCACC	407
Db	111	ITLVLSELSERSESRPRLAIVALLIENSRERVALSERALILESERLEUPROTHR	130
QY	408	ACCAATGTCAGGCCGAGGCACTGTCTGTCTACTCTCAGGTTTGGACTGGAGCCAGAAAAC	467
Db	131	ALAProProlIALAIGLYTHrcIUbsrleuIleSerGIYTrpGLYAsnThrIeuser	150
QY	468	AGTGGCCCAACACCTGCACTTGGCGAGAACTGGAGGCCCCGTATGTCTGATCGAGA	527
Db	151	GLYAlaAepTYrProAspGIuIeueGInCYsIleuAspAlaPProValIeuserGInaIgu	170
QY	528	TGCCAAAAAACGACAAAGAAAGAAAGCCACAGACATCTCTTATGTGTGAAATTGTGAA	587
Db	171	CysGInAlaIasErTYrProGIYLYsIleThrAsnAsnMetPheCYsValGIYpHeIeugIu	190
QY	588	GTATTACACCCGAAATTTTGGGAGGCGCCGTGTGCTACTGTCACTGTCAAGCAAGCTC	647
Db	191	GLYGLYLYsAepSerCYsGInGLYsAepSerGLYGLYProValIValSerAsnGLYIuEun	210
QY	648	CAGGGAATC-----GAGGTGGGGCACTTCATGGAGGAGGAGCGTGGCACTACACC	698
Db	211	GIingIYIleValSerTrpGLYTYrGLYCYsAlaGInLYsAsnAryProGLYValTYrThr	230
QY	699	AATGTTTCAAAATATGTATCCGTGATGGAAACT	734
Db	231	LYsValTYrAsnTYrValAspTrpIleLYsAepThr	242

RESULT 14

TRY2_XENLA

ID	TRY2_XENLA	STANDARD:	PRT;	244	AA.
AC	P70059;				
DT	01-NOV-1997	(Rel. 35, Created)			
DT	01-NOV-1997	(Rel. 35, Last sequence update)			
DE	28-FEB-2003	(Rel. 41, Last annotation update)			
DT	Trypsin precursor	(EC 3.4.21.4)			
OS	Xenopus laevis (African clawed frog).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;				
OC	Xenopodinae; Xenopus.				
OX	NCBI_TaxID=8335;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Wang K., Lytle L., Gan L., Hood L.E.;				
RL	Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.				
CC	-1- CATALYTIC ACTIVITY: Preferential cleavage: Arg -Xaa, Lys -Xaa.				
CC	-1- SUBCELLULAR LOCATION: Extracellular				
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation-				
CC	at the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/)				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL: U72330; AAB17274.1; -.				
DR	HSSP: P00763; IDBO.				
DR	MEROPS: S01.258; -.				
DR	InterPro: IPR001314; Chymotrypsin.				
DR	InterPro: IPR001254; Ser_protease_Try.				
DR	Pfam: PF00089; trypsin; 1.				
DR	PRINTS: PR00722; CHYMOTRYPSIN.				
DR	SMART: SM00020; Tryp_Spc; 1.				
DR	PROSITE: PS0240; TRYPsin_DOM; 1.				
DR	PROSITE: PS00134; TRYPsin_HIS; 1.				
DR	PROSITE: PS00135; TRYPsin_SER; 1.				
KW	Hydrolase; Serine protease; Digestion; Zymogen; Signal;				
KW	Multigene family.				
FT	SIGNAL	1	15	BY SIMILARITY.	
FT	PROPEP	16	21	ACTIVATION PEPTIDE (BY SIMILARITY).	
FT	CHAIN	22	244	TRYPsin.	

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 8, 2003, 00:24:46 ; Search time 72 Seconds
(without alignments)
6207.599 Million cell updates/sec

Title: US-10-037-270-482
Perfect score: 1560
Sequence: 1 ggcactactccctcagcta.....cggagatgtagcatcagt 866

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 830525 segs, 258052604 residues
Total number of hits satisfying chosen parameters: 1661050

Minimum DB seg length: 0
Maximum DB seg length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-O=fgn2.1/USPTO_spool/US10037270/runat_07112003.140515.8759/app_query.fasta_1.10311
-DB=SPTRMBL_23 -QMT=faeacan -SUFFIX=irpt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=-1 -END=-1 -MATRIX=blomsun62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10037270.GCGN.1.1.71.@runat_07112003.140515.8759 -NCPUB=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: SPTRMBL_23.*
2: sp_archaea.*
3: sp_bacteria.*
4: sp_fungi.*
5: sp_human.*
6: sp_invertebrate.*
7: sp_mammal.*
8: sp_mmc.*
9: sp_organelle.*
10: sp_phase.*
11: sp_plant.*
12: sp_rodent.*
13: sp_virus.*
14: sp_vertebrate.*
15: sp_unclassified.*
16: sp_rv1rus.*
17: sp_bacteriopl.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Length	DB ID	Description
1	1065	68.3	237	11 Q9DAA4 Q9daa4 mus musculu

2	483	31.0	112	11 Q9CUH3 Q9cuh3 mus musculu
3	347.5	22.3	246	11 Q9ROT7 Q9rot7 mus musculu
4	347.5	22.3	251	4 Q8N2U3 Q8n2u3 homo sapien
5	341.5	21.9	246	11 Q9QUX9 Q9qux9 mus musculu
6	331	21.2	247	4 Q8NWM4 Q8nwm4 homo sapien
7	328.5	21.1	246	11 Q9Z1R9 Q9z1r9 mus musculu
8	328.5	21.1	247	11 Q9D7Y7 Q9d7y7 mus musculu
9	328.5	21.1	247	11 Q9CPN9 Q9cpn9 mus musculu
10	327.5	21.0	239	4 Q8N1C9 Q8n1c9 mus musculu
11	315	20.2	247	11 Q9CPN7 Q9cpn7 mus musculu
12	311	19.9	242	13 Q9W7Q7 Q9w7q7 paralicthy
13	305.5	19.6	238	13 Q9W7Q6 Q9w7q6 paralicthy
14	305	19.6	241	4 Q81YP2 Q81yp2 homo sapien
15	299.5	19.2	244	13 Q98TG9 Q98tg9 engraulis j
16	299.5	19.2	244	13 Q42159 Q42159 petromyzon
17	299	19.2	237	13 Q91515 Q91515 fugu rubrip
18	297.5	19.1	245	13 Q42160 Q42160 petromyzon
19	290.5	18.6	240	13 Q98TH0 Q98th0 engraulis j
20	290	18.6	242	13 Q93266 Q93266 pseudopleur
21	287	18.4	242	13 Q92099 Q92099 paraneotothe
22	284	18.2	241	11 Q8BW11 Q8bw11 mus musculu
23	282.5	18.1	247	13 Q42158 Q42158 petromyzon
24	282.5	18.1	247	13 Q42608 Q42608 petromyzon
25	282	18.1	222	13 Q8AV11 Q8av11 oncorhynch
26	280	17.9	255	4 Q96RQ0 Q96rq0 homo sapien
27	276.5	17.7	243	13 Q8AVB3 Q8avb3 brachydario
28	267	17.1	219	13 Q91036 Q91036 gadus morhu
29	263	16.9	244	13 Q8QW3 Q8qw3 anguilla ja
30	261.5	16.8	251	11 Q9D9G7 Q9d9g7 mus musculu
31	260.5	16.7	251	11 Q54854 Q54854 rattus norv
32	255.5	16.4	253	11 Q91Y82 Q91y82 mus musculu
33	253.5	16.2	178	13 Q93594 Q93594 dicentrarch
34	253.5	16.2	250	13 Q93265 Q93265 pseudopleur
35	252.5	16.2	246	11 Q88301 Q88301 mus musculu
36	249	16.0	260	4 Q81W69 Q81w69 homo sapien
37	247	15.8	249	13 Q92046 Q92046 dissostichu
38	244	15.6	293	11 Q9D140 Q9d140 mus musculu
39	242	15.5	249	13 Q9W6K0 Q9w6k0 notochtha
40	242	15.5	249	13 Q81PQ7 Q81pq7 gadus morhu
41	240	15.4	675	13 Q9W6J8 Q9w6j8 dissostichu
42	238.5	15.3	249	11 Q91VE3 Q91ve3 mus musculu
43	238	15.3	344	13 Q9W6J9 Q9w6j9 dissostichu
44	237	15.2	248	11 Q63224 Q63224 rattus norv
45	233	14.9	255	11 Q9J182 Q9j182 mus musculu

ALIGNMENTS

RESULT 1
ID Q9DAA4 PRELIMINARY; PRT: 237 AA.
AC Q9DAA4;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE 1700016G0SRik protein (Hypothetical serine proteases).
GN 1700016G0SRik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa T., Hara A., Fukunishi Y., Komno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamane A. I.,
RA Saito T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H. A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsumoto Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L. M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bulc C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hornann M., Humé D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Maaham M., Mazzarelli U., Mombearts P.,
RA Nordone P., Ring B., Schenckel M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Ringschwald C., Seya T., Shibaie Y., Storch K.-F.,
RA Suzuki H., Toyoo-oka K., Wang K.H., Wetzl C., Wittkeker C., Witting L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 409:685-690(2001).
RN [2]

SEQUENCE FROM N.A.
RP
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=32354683; PubMed=12466851;
RA THE FANTOM Consortium,
RA the Riken Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL: AK006028; BAB24373.1; -.
DR EMBL: P007980; BAC26613.1; -.
DR HSSP: PO0761; IAKS.
DR MEROPS: S01_989; -.
DR MGd: MG11914940; 1700016G05R1K.
DR Interpro: IPR001314; Chymotrypsin.
DR Interpro: IPR001254; Ser protease_Try.
DR Pfam: PF00089; trypsin_1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; TRYD_SPC_1.
DR PROSITE: PSS0240; TRYPsin_DOM_1.
DR PROSITE: PSS0240; TRYPsin_DOM_1.
DR Hypothetical protein; Hydrolase; Protease; Serine protease.
QO SSQUENCE 237 AA; 2669 MW; EF9139656B0394A CR664;

	Alignment Scores:			
Pred. No.:	2, 06e-101	Length:	237	
Score:	1065.00	Matches:	196	
Percent Similarity:	90.64%	Conservative:	17	
Best Local Similarity:	83.40%	Mismatches:	22	
Query Match:	68.27%	Indels:	0	
DB:	11	Gaps:	0	
US-10-037-270-482 (1-866) x Q9DA44 (1-237)				
QY	42 ATGAAATATGTCTTATTATTTGGAGTCSTCGTGGAACATTTTTCTTGACTCATCT	101		
Db	1 MetLysLeuIleIerLytleuthkrlieuaIaglytlrAlaleuValthrHISserSer	20		
QY	102 GTTCAGAAAGAACCCCTGCCTCCTATTTGGTGNACSTCAAGTCTCAACCCCTGT	161		
Db	21 ValGIILyIselIuAsprHISalAProLytleuaIatLytleuLyserrPhaeAsPrCoS	40		
QY	162 GTGGGGTCTCTATCAAACCCAGTCGGGNGTGGGCCACGTCATCTGTATTTCAAAT	221		
Db	41 ValGIValleuIleLysalaserTrpaValleuaIarPseSerHIScyTYLecraSn	60		
QY	222 CTGAAGAATGATCTGGGAAATTTCAAGACACAGTCAAGACAGGATCTGAACAACAATT	281		
Db	61 LeuarGyalMetleuclIyasnRheuyssrraIvalaIdgarBgLytlrGlugInThrIle	80		
QY	282 AACCCCATTCAGATCTGCCGCTACTGGAACTAACATCACTCAATAGCGCCCAACAGATGACTC	341		
Db	81 TytProlledInIlerIearGtylrTraenTyUserHISethrAlaProGIlnAsprIeu	100		
QY	342 ATGCTCATCAAGTGGCTAAGCTGGCATCTCAATCCCAAGTCAAGCCCTTCCTC	401		
Db	101 MetleuIleLysleuaIalysProlatrRheaenHISlysValGIlnalleuProile	120		
QY	402 GGCACACCAATCTCAAGGCACAGCACGTGTCATCTCACTCAAGCTTGGACCTGAGGCA	461		
Db	121 AlaThrThrasnaIarGrProGIlnrValCythrIleuSeGlyLeuaSPrtSerIn	140		

QY	462	GAAGAAAGGCGGACACCCCTGACCTTGCGGAGAACTGGAGGCGCCCGGTGAGTCTGAT	520
	141	GIUAGAGAGClyAyrHieProapehLeuArgGlnAsnLeuGluAlaProValMetThrAsp	160
QY	522	CGAGAAATGCGAAAAACAGAACAGAAAAAGCCACAGAAATTCCTTATGTGTGAATTT	581
	161	LyAspCyGlnLyThrGlnGlnGlySerSerHisArgAsnSerLeuCyValArgPhe	180
QY	582	GTGAAAGTATTCAGCGCAATTTTGGGAGAGTGGCCGCTGTCATCTGCATCTCGAAAGAC	641
	181	ValLySValPheSerArgGilepheGlyGluValAlaValAlaThrValIleCyLysAsn	200
QY	642	AAGCTCCAGAGAAATGAGAGTGGGCGCACTTCATGGAGAGGGGACGTCGGCATCTACACCAAT	701
	201	LyLeuGlnGlnGlyIleGluValAlGlyHisIlehetecIyGlyAspValGlyIleTyThrAsn	220
QY	702	GTTTACAAATATGTATCTGATTCGATTCAGAAACCTGCTAAGCAAG 746	
	221	IleTySerTyValProTyrIleGlyIlySerThrThrThrSerGlyIlyS 235	
RESULT 2			
Q9CCH3	PRELIMINARY;	PRT;	112 AA.
AC	Q9CCH3;		
DT	01-JUN-2001 (TREMBLrel. 17, Created)		
DT	01-JUN-2001 (T-EMBLrel. 17, Last sequence update)		
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)		
DE	1700016GOSRIK protein (Fragment).		
GN	1700016GOSRIK.		
OS	Mus musculus (Mouse).		
OC	Mammalia; Eutheria; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxId=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Testis;		
RX	MEDLINE=21085660; PubMed=11217851;		
RA	Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ihii Y.,		
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,		
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,		
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,		
RA	Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,		
RA	Pfeischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,		
RA	Kuehl P., Lewis S., Matsuo Y., Nikaizo I., Pesole G., Quackenbush J.,		
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,		
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barz G.,		
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,		
RA	Brumstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,		
RA	Gustringich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,		
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,		
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,		
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,		
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,		
RA	Yusufaw-Boris A., Yoshida K., Haesgawa Y., Kawai J.H., Kontsuki S.,		
RA	Hayashizaki Y.;		
RT	"Functional annotation of a full-length mouse cDNA collection.";		
RL	Nature 409:685-690(2001).		
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY SL.		
DR	EMBL; AK016088; BAB30109.1; -.		
DR	HSSP; P00761; IAKS.		
DR	MEROPS; S01.989; -.		
DR	MGD; MG1:1914940; 1700016GOSRIK.		
DR	InterPro; IPR001314; Chymotrypsin.		
DR	InterPro; IPR001254; Ser. protease_Try.		
DR	pfam; PF00689; trypsin; 1.		
DR	PRINTS; PR00722; CHYMOTRYPsin.		
DR	Hydrolase; Protease; Serine protease.		
FT	NON TER		
FT	112 112		
SQ	SEQUENCE	112 AA; 12764 MW; B95A276D81352923 CRC64;	


```

QY 588 GTATTGACCGCAATTTTGGGAGGTGGCCGTGCTACTGTCATCTGCAAGAACAGCTC 647
Db 191 GtGlyLysAspSerCysGlnGlyAspSerGlyGlyProValValCysAsnGlyGlnLeu 210
QY 648 CAGGGAATC-----GAGTGGGGGACCTTCATGAGGAGGAGCGTGGCATCTACAC 698
Db 211 GtGlyLysValSerTrpGlyTrpGlyCysAlaLeuLysAspAsnProGlyValTyrThr 230
QY 699 AATGTTACAAATATGATCTGATTCGATTCAGACACT 734
Db 231 LysValCysAsnTyrValAspTrpIleGlnAsnThr 242

RESULT 4
Q8N2U3 PRELIMINARY; PRT; 251 AA.
ID Q8N2U3;
AC Q8N2U3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Strauberg R.;
Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; BC030238; AAH30238.1;
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser. protease_Try.
DR Pfam; PF00069; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS50240; TRYPsin DOM; 1.
DR PROSITE; PS00134; TRYPsin HIS; 1.
KM Hypothetical protein; Hydrolase; Protease; Serine protease.
FT NON_TER
SQ SEQUENCE 251 AA; 27091 MW; 53D3993B585AB328 CRC64;

Alignment Scores:
Pred. No.: 5 28e-27 Length: 251
Score: 347.50 Matches: 82
Percent Similarity: 52.59% Conservative: 40
Best Local Similarity: 35.34% Mismatches: 103
Query Match: 22.28% Indels: 7
Gaps: 3

US-10-037-270-482 (1-866) x Q8N2U3 (1-251)
QY 60 TTGGGTGCTCGCTGGGACATTTTCTTGCTACTCATCTGTCAG----- 107
Db 15 LcGlyLThrValAlaValProPheAspAspAspAspLysLysLysLysLysLysLys 34
QY 108 AAGAAGACCCCTCTCTCTTTTGGTACTCAAGTCTCACTTCAACCCCTGTGGGC 167
Db 35 GtGlnAsnSerLeuProTyrGlnValSerLeuAsnSerLysSerHisPheCysGlyGly 54
QY 168 GTCCCTCAACCAACCGCTGGTGGGCGCCGCTACGCTCATCTTACCAATTCGAAA 227
Db 55 SerLeuIleSerGlnGlnTrpValValSerAlaAlaHisCysTyrLysThrArgIleGln 74
QY 228 GTGATCTGGGAATTTCAAGACAGAGTCAAGCGTACTGACAGACAGACATTAACCC 287
Db 75 ValArgLcGlyLysLysAsnLysValLeuGlnGlyAsnGlnGlnPheIleAsnAla 94
QY 288 ATTCAAGTCTCGGCTACTGGAATCAAGTCAAGTCAAGCGCCCAAGATGACCTCATGCTC 347
Db 95 AlaLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 114

```

```

QY 348 ATCAAGCTGAGCTAAGCTGCGCATCTCAATCCCAAGTCAGCCCTTCCCTCGCCACC 407
Db 115 IleLysLeuSerSerProAlaValIleAsnAlaArgValSerThrIleSerLeuProThr 134
QY 408 ACCAATGTCAGCCAGGACCTGCTCTTACTTCACTGAGTTTGATCGAGCCCAAGAAC 467
Db 135 AlaProProAlaAlaIleGlyThrGlyCysLeuIleSerGlyTyrpLysAnthrLeuSerPhe 154
QY 468 AGTGGCGGACCCCTGACTTGGCGGAGAACCTGAGAGCCCGCGTGAATGTCAGGAA 527
Db 155 GlyAlaAspTyrTrpAspGlyLeuLysCysLeuAspAlaProValLeuThrGlnAlaGln 174
QY 528 TGCCAAAAAACAACAGAGAAAGAAAGCAAGCAAGATCTTATGTCGAATTTGAAA 587
Db 175 CysLeuAlaSerTyrTrpAspGlyLeuLysCysLeuAspAlaProValLeuThrGlnAlaGln 194
QY 588 GTATTGACCGCAATTTTGGGAGGTGGCCGTGCTACTGTCATCTGCAAGAACAGCTC 647
Db 195 GtGlyLysAspSerCysGlnArgAspSerGlyGlyProValValCysAsnGlyGlnLeu 214
QY 648 CAGGGA---ATCGAGGTGGGGACCTTCATGGGA-----GGGAGCTGGGCATCTACAC 698
Db 215 GtGlyLysValSerTrpGlyHisGlyCysAlaTrpLysAsnArgProGlyValTyrThr 234
QY 699 AATGTTACAAATATGATCTGATTCGATTCAGACACT 734
Db 235 LysValTyrAsnTyrValAspTrpIleLysAspThr 246

RESULT 5
Q9OUK9 PRELIMINARY; PRT; 246 AA.
ID Q9OUK9;
AC Q9OUK9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE TESP4 (0910001B1PRK OR TRYPsinogen 9).
GN TC OR 0910001B1PRK OR TRYPsinogen.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA MEDLINE=99436155; PubMed=10506205;
RX Kashiwabara S., Baba T.;
RA Ohmura K., Kohno N., Kobayashi Y., Yamagata K., Sato S.,
" A homologue of pancreatic trypsin is localized in the acrosome of
mammalian sperm and is released during acrosome reaction.";
RL J. Biol. Chem. 274:29426-29432(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Pancreas;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada I.,
Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barch G.,
Blake U., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohseki S.,
Hayashizaki Y.;
"Functional annotation of a full-length mouse cDNA collection.";

```


Db	29	ThrcygluclulsmbsevalProtyrGlnIvalserleuamSerclySerhiserPheCys	48		
Qy	162	GTGGGCGCTCTCTCAAAACCCAGCTGGGTGGTGGCCCCCACTACTGCTATTATCAAAAT	221		
Db	49	GlyGlyserleuileSerGlnIntPrValIvalSerAlaGlnIleCysGlyLeuProHis	68		
Qy	222	CTGAAAGTGAATGCTGGGAAATTTCAAGACACAGTCAGAGACCGATCTGAACAGAAATT	281		
Db	69	IleGlnIvalArgLeuGlyGlnHisIleValIleGluValLeuGlnGlyAenGlnGlnPheIle	88		
Qy	282	AACCCCATTCAGATCTCGCTGCTACTGGAACTACAGTCATAGCGCCSCACAGATGACTC	341		
Db	89	AsnAlaAlaIleValIleLeuArgHisProGlyrTyraAsnArgIleThrLeuAsnAsnAspIle	108		
Qy	342	ATGCTCATCAAGCTGCTAAGCCCTGATGCTCAATGCCAAAGTCCAGCCCTTCCCTC	401		
Db	109	MetLeuIleLeuSerThrProAlaValIleAsnAlaHisIleValSerThrIleSerLeu	128		
Qy	402	GCCACACCAATGCTACAGCCGACGACATGTCTGTACTCTGACGTTTGAGACTGGAGCCAA	461		
Db	129	ProThrIalProProAlaIalAglIthrGlnCysLeuIleIleSerGlyItrpGlyAsnThrIleu	148		
Qy	462	GAAGACAGTGGCGGACACCCCTGACTTGGCGAGAACTGGAGAGCCCGCTGATGCTGAT	521		
Db	149	SerSerGlyAlaAspIlyrProAspGluLeuGlnCysLeuAsnAlaProValIleuThrGln	168		
Qy	522	CGAAGATGCCAAAAACAGAACAGGAAAAACCCACAGAAATTCCTTAATGTGTGAATTT	581		
Db	169	AlaIleCysAlaIleSerThrProLeuAlaValIleThrSerLeuMetPheCysValGlyPhe	188		
Qy	582	GTCGAAAGTATTCAGCCGGAATTTTGGGGAGGTGGCCGCTTGCTACTGTATCTGCAGAAAC	641		
Db	189	LeuGlnGlyGlyIlyAsnAspSerCysGlnGlyIlyAsnSerGlyIlyProValIalCysAsnGly	208		
Qy	642	AGGCTCCAGGGATTC-----GAGGTGGGGCACTTCATGAGGAGGGAGGTGGGATC	692		
Db	209	GlnLeuGlnIleValIleValSerTrpGlyrGlyCysAlaGlnIlyAsnArgItrpGlyVal	228		
Qy	693	TACACCAATGTTTACAAATATGATATCTCGATTCGATTCGAAACACT	724		
Db	229	TyrThrIlyValIlyAsnIlyrValIalTrpIleIlyAsnTrpThr	242		
RESULT 7					
Q9Z1R9	ID	PRELIMINARY;	PRT; 246 AA.		
AC	Q9Z1R9				
DT	01-MAY-1999	(TEMBLRel. 10, Created)			
DT	01-MAY-1999	(TEMBLRel. 10, Last sequence update)			
DT	01-MAY-2003	(TEMBLRel. 23, Last annotation update)			
DE	Trypsinogen 16.				
GN	TRYGN16 OR TRYPsinOGEN.				
OS	Mus musculus (Mouse).				
OC	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Euteheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_faxid=10090;				
RN	1	SEQUENCE FROM N.A.			
RP	1	SEQUENCE FROM N.A.			
RC	STRAIN-BALB/c;				
RA	Rowen L.; Hood L.;				
RT	"Comparison between strains Balb/C and 129 in a region of the mouse T				
RT	cell receptor beta locus."				
RL	Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.				
RN	12				
RC	STRAIN-BALB/c;				
RA	MEDLINE=21103195; PubMed=11160223;				
RA	Chen F., Rowen L., Hood L., Rothenberg E.V.;				
RT	"Differential transcriptional regulation of individual TCR Vbeta				
RT	segments before gene rearrangement."				
RL	J. Immunol. 166:1771-1780(2001)				
CC	- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.				
OR	EMBL; AF107342; AAC79093.1; -.				

[illegible]

DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE 2210010C04Rik protein.
 GN 2210010C04Rik.
 OS Mus musculus (Mouse).
 OC Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Stomach;
 RX MEDLINE=21085660; PubMed=11217651;
 RA Kawai T., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Atakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa K., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gijobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Glisic C., King B., Kochia H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirkl L.M., Staudt P., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamlya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli U., Nombarts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whiteaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsaki S.,
 RA Hayashizaki Y.
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 DR EMBL; AK008695; BAB25837.1; -
 DR HSSP; P00763; IDPO.
 DR MEROPS; S01.151; -
 DR MGD; MGI:1914623; 2210010C04Rik.
 DR Interpro; IPR001314; Chymotrypsin.
 DR Interpro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PSS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR KEGG; 04310; trypsin; Serine protease.
 SO SEQUENCE 247 AA; 26407 MW; 84d474db322b3a55 CRC64;

Alignment Scores:
 Pred. No.: 4.9e-25 Length: 247
 Score: 328.50 Matches: 75
 Percent Similarity: 53.14% Conservative: 35
 Best Local Similarity: 36.23% Mismatches: 94
 Query Match: 21.06% Indels: 3
 DB: Gaps: 1

US-10-037-270-482 (1-866) x Q9D7Y7 (1-247)

QY	123	CCCTATTGTGTGTA	CTCAAGTCTCA	ACCCCTGTGGGCTCTT	CTATCAAA	CCC	182
DB	37	PrIotyGlnValSer	LeuAsnSerGly	TrHisPheCyGly	GlySerLeuIleAsnSer		56
QY	183	AGCGGGGTGTGG	CCCCAGCTCA	CGTCTATT	TTCACAAATTCGA	AGATGCTGGGAAT	242
DB	57	GlnTrpValValSer	IleAlaHisCyTrp	ValSerAla	GlnValAla	ArgLeuGlyGlu	76
QY	243	TTCAGAGCGAG	CTCAGAGCGG	TACTGGA	CAGACATTAAC	CCCATTCAGATCGTCGC	302
DB	77	HisAsnIleAsp	AlaLeuGlnGly	GlnGlnInp	HeIleAspAlaAla	ValyIleIleArg	96
QY	303	TACTGAACTCA	AGTCAATAG	CGCCCA	AGATGAC	TACCTCTATCA	CTAGCGTTAAG

Db 97 HsiproantnyraenAlaAntThrtYranhsnAepIIEmetLeuIIeYsLeuYsthr 111
 Qy 363 CCTGCCATGCTCAATCCCAAGTCAGGCCCTTCCCTCGGCACCAACCATGTCAAGCCA 422
 Db 117 AlAlaThrLeuAanserArgValserThrValAlaLeuProArgserCysPProserAla 136
 Qy 423 GGCACTGTCTGTCTACTCTCAAGCTTTGGACTCGAGCCCAAGAAACAGTGGCCGACACCT 482
 Db 137 GLythrArgCysLeuValserGclYTrpGlyAntntrLeuSerSerGclYThrAntyPro 156
 Qy 483 GACTTGGCGGCAACCTCGAGAGCCGCCCTGTAGTGTGATTCGAGTAATCCAAAAACAGAA 542
 Db 157 SerLeuLeuGlnCysLeuAspAlaProValLeuSerAspSerCysThrserSerTy 176
 Qy 543 CAAGCAAAAAGCCACAGGAATTCCTTAGTGTGAATTTGTGAAGATTCATGAGCCGAATT 602
 Db 177 ProGlytyrGlyLeuThrSerAsnMetPheCysLeuGlnGlyLeuGlnGlyAspSer 196
 Qy 603 TTTGGGAGAGTGGCCGCTTGCTACTGTCTATCTGCAAAAGCAACACTCAGGAGATC----- 656
 Db 197 CysGlnGlyAspSerGlyGlyProValValCysAsnGlyGlnLeuGlnIlyValValSer 216
 Qy 657 ---GAGNGGGGCGCACTTCATGGAGGGGAGCGCGGCACTTACCAACCAATTTTCAATAT 713
 Db 217 TrpGlytyrGlyCysAlaGlnArgGlySerProGlyValTyThrIlyValCysIyTy 236
 Qy 714 GTATCTCGATGTGAGAACT 734
 Db 237 ValAsnTrpIleGlnGlnThr 243
 RESULT 9
 Q9CPN9 PRELIMINARY; PRT; 247 AA.
 AC Q9CPN9;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE 2210010C04RIK protein (Trypsinogen 7).
 GN 2210010C04RIK OR TRYPSINOGEN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6J; TISSUE=Small intestine, and Pancreas;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinnagawa A., Shibata K., Yoshino M., Itch M., Ishii Y.,
 Arikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 Aikawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamamaki I.,
 Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 Kadota K., Matsuda H.A., Aeshburner M., Batalov S., Casavant T.,
 Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,
 Kuhl P., Lewis S., Matsuo Y., Nkaido I., Pesole G., Quackenbush J.,
 Schriml L.M., Staibli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 Rosenblum M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 Guenrich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,
 Lyons P., Marchionni L., Maehima J., Mazzarelli I., Sakamoto N.,
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 Saeki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 Sasaki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 Wynshaw-Boris A., Yoshida K., Haegawa Y., Kawaji H., Kohzuki S.,
 Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:665-690(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c;
 RX MEDLINE=21103195; PubMed=11160223;
 RT Chen F., Rowen L., Hood L., Rothenberg E.V.,
 "Differential transcriptional regulation of individual TCR Vbeta

QY 648 CAGGAGATC-----GAGGTGGGGACCTTCATGGAGGGAGCGTGGCATTCACCC 668
 |||||
 Db 211 GlnGlyIleIvalSerTrpGlyIyrGlyCysAlaGlnIyrAsnAngProGlyValIyrThr 230
 |||||
 QY 699 AATGTTTACAAATATGATCTCGATT 725
 |||||
 Db 231 LyValIyrAsnIyrValAspTrpIle 239
 |||||
 RESULT 11
 Q9C9N7 PRELIMINARY; PRT; 247 AA.
 AC Q9C9N7;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE 1810009J06Rik protein (Trypsinogen 4).
 GN 1810009J06Rik OR TRYPSINOGEN.
 OS Mus musculus (Mouse).
 OC Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Pancreas;
 RX MEDLINE=21085660; PubMed=1217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Izawa A., Fukunishi Y., Kono H., Aichi U., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Glass C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Maehiro T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Balarelli R., Barsh G.,
 RA Blake J., Bonfelli D., Bojunga C., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.F.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaserts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyokawa K., Wang K.H., Weltz C., Whitlaker C., Wilming L.,
 RA Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohetsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RN Nucleic Acids Res. 29(12):4099-4103 (2001).
 RP Nucleotide 409:685-690(2001).
 RA [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c;
 RX MEDLINE=21103195; PubMed=11160223;
 RA Chen F., Rowen L., Hood L., Rothenberg E.V.;
 RT "Differential transcriptional regulation of individual TCR Vbeta
 RT segments before gene rearrangement.";
 RN J. Immunol. 166:1771-1780(2001).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 DR EMBL; AK007406; BAB5018.1; -;
 DR EMBL; AE000663; AAB69055.1; -;
 DR HSSP; P00763; IDPO.
 DR MEROPS: S01.129; -;
 DR MGD; MG1192087; 1810009J06Rik.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00723; CHYMOTRYPSIN.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PSS0240; TRYPsin_DOM; 1.
 DR PROSITE; PS00134; TRYPsin_HIS; 1.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 KM Hydrolase; Protease; Serine protease.
 SQ SEQUENCE 247 AA; 26503 MW; ED5F6796833C7BC CRC64;

Score:	315.00	Matches:	81
Percent Similarity:	53.27%	Conservative:	33
Best Local Similarity:	37.85%	Mismatches:	84
Query Match:	20.19%	Indels:	16
DB:	11	Gaps:	6

US-10-037-270-482 (1-866) x Q9CPN7 (1-247)

```

QY      123 CCTATTGGTGTACTCAG-----TCGATCTCAACCCCTGTGGCGTCTC 173
           |||||
Db       36 ProTyglnValSerLeuAAsrPrLyIleSerHis-----GInCyGlyGlySerLeu 53
           |||||
QY      174 ATCAAAACCAGCTGGGTGTGGCCCCACCTACTACGTATTATCAAAATCGAAATGATG 233
           |||||
Db       54 IleSerArgGlnTrpValLeuSerLAlaHisCyGlyTrpValGdAglyGlnValArg 73
           |||||
QY      234 CTGGGAATTTTCAGAAGCAGATCAGAACCGTGACTGAAACAGAAATAAACCCCATTCAG 293
           |||||
Db       74 LeuGlYgLnHlaAsnIlLaArValLeuSluGlYglGlnGlnPheIlLaArAlaGluLe 93
           |||||
QY      294 ATGTGTCGGTCTGAGTAACAATCAGTATGACGCCCCCAAGATGACSTCATGCTCATGAG 353
           |||||
Db       94 IleIleArgHisProAAsrPrLyTrpValAsrThrValAAsrAAsrIlMeLeuIlLe 113
           |||||
QY      354 CTGGCGTAACCGCCGATGCTGATCCCAAACTCCAG-----CCCGTCCCGCTC 402
           |||||
Db      114 LeuLYSerProAlaIlLeuAAsrSerTrpValSerLeuProAAsrSerCys 133
           |||||
QY      402 GCCACACCAATATGATAGGCAGGACACTGTGTGTACTCTAGATTTGACATGGAGCAA 461
           |||||
Db     134 AlaSerThrAsnAlaGln-----CysLeuValSerGlyTrpGlyAAsrThrVal 149
           |||||
QY      462 GAAAACAGTGGCCGACCCCTGACTTGGCGAGAACTGGAGGCCCGCTGATGTGAT 521
           |||||
Db     150 SerIleGlyGlyTrpProAlaLeuLeuGlnGlySerGlnAlaProValLeuSerAla 169
           |||||
QY      522 CGAAGATGCCAAAAACAGAACAGAAAGAAACCAAGAAATTCCTATGTGTGAATTT 581
           |||||
Db     170 SerSerCysValyAsrSerTrpProGlyGlnIleThrSerAsnMetPheCysLeuGlyPhe 189
           |||||
QY      582 GTGAAGATTTTCAGCGCAATTTTGGGAGGTGGCCGTGTACTGTACATCTGCAAGAC 641
           |||||
Db     190 LeuGlnGlyGlyLeuAAsrSerCysAAsrPrLyAAsrSerGlyProValValCysAAsrGly 209
           |||||
QY      642 AAGCTCCAGGCAATC---GAGGTGGGCACTTC-----ATGGAGGCGGACGTGGCATC 692
           |||||
Db     210 GluIleGlnGlyIleValSerTrpGlySerValCysAlaMetAAsrGlyLeuProGlyVal 229
           |||||
QY      693 TACACCAATGTTTACAAATATGATATCTGCGATTGAGAACT 734
           |||||
Db     230 TyrThrLYValCYAAsrTrpLeuSerTrpIleGlnIuThr 243
           |||||

```

RESULT 12
Q9W707 PRELIMINARY; PRT: 242 AA.

AC Q9W707;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
OC Actinopterygii; Neopterygii; Teleostei; Euteleostomi;
OC Acanthomorpha; Acanthopterygii; Percomorphia; Pleuronectiformes;
OC Pleuronectidae; Paralichthyidae; Paralichthys.
NCBI_TaxId=2255;
RN NCB1
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RA Suzuki T., Suiyastrava A.S., Kurokawa T.;
RT "Japanese flounder mRNA for tryptogenin 1.",
RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases:
EMBL; AB029750; BAA82362.1; -

DR HSP; P00763; IDPO.
 DR MEROPS; S01.125; -.
 DR InterPro; IPRO01254; Ser protease_Try.
 DR Pfam; PF00089; trypsin_1.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS50240; TRYP_SIN_DOM; 1.
 DR PROSITE; PS00134; TRYP_SIN_HIS; 1.
 DR PROSITE; PS00135; TRYP_SIN_SER; 1.
 DR Hydrolyase; Protease; Serine protease.
 SQ SEQUENCE 242 AA; 26548 MW; 6DA722C80BC194A2 CRC64;

Alignment Scores:

Pred. No.:	3.18e-23	Length:	242
Score:	311.00	Matches:	69
Percent Similarity:	53.52%	Conservative:	45
Best Local Similarity:	32.39%	Mismatches:	83
Query Match:	19.94%	Indels:	16
DB:	13	Gaps:	4

US-10-037-270-482 (1-866) x Q9W7Q7 (1-242)

QY 123 CCTATTGGTGAACCTCAAGTCTCACTCAACCCCTGTGGGCGTTCATCAACC 182
 DB 33 ProhibitinValSerLeuAsnSerGlyTyrHisPheCysGlySerLeuValaAnglu 52
 QY 183 AGCTGGGTGGTGGCCCGCCAGCTCACTGCTATTACCAATGGAAGTATGCTGGAAAT 242
 DB 53 AsnTrpValSerAlaAlaHisCysTyrLysSerArgValaGluValaGluMetGlyglu 72
 QY 243 TTCAGAGCAGAGTCAGACGCGTACTGAAACAGCAATTAACCCCATTCAGATGCTCGC 302
 DB 73 HisHisIleLysIleAsnGluGlyThrGluGlnPheIleSerSerGluArgValaIlearg 92
 QY 303 TACTGGAATCACTGATAGCGCCCAAGAGTACCTGATGCTCATCAAGTGGCTAAG 362
 DB 93 HisProAsnTyrSerSerTyrAsnIleAsnAsnAspIleMetLeuIleLysLeuArglu 112
 QY 363 CCTGCCATGCTCATCCCAAGTCCAGCCCTTCCCTCCGCAACCAATGTCAGGCCA 422
 DB 113 ProIleThrLeuAsnGlnTyrValaGlnProValaIleLeuProThrSerCysAlaProIle 132
 QY 423 GGCATGCTGTCTACTCTCAAGTGGTGGATGAGCCAGCAAGAA--AAAGTGGCCGACAC 479
 DB 133 GlyThrMetCysThrValSerGly-----TrpGlyAsnThrMetSerSerThrAlaAsn 150
 QY 480 CCGACTGTGGGCAAGCTGAGAGGCCCGGTATGCTGTAATTTGGAAGTATTCAGCCGA 539
 DB 151 ArgAspMetLeuGlnCysLeuAspLeuProIleLeuSerAspArgAspCysGluAsnSer 170
 QY 540 GAACAAAGAAAAGCCACAGAAATTCCTTATGTGTAATTTGGAAGTATTCAGCCGA 599
 DB 171 TyrProIleMetIleThrProAlaMetPheCysAlaGlyTyrLeuGluGlyLysAsp 190
 QY 600 ATTTTGGGAGGTGGCGCTGTACTGCTCATCTGCAAGAACAAGTCCAGGGAAATC-- 656
 DB 191 SerCysGlnGlnLysSerArgLysGlyProValaValCysAsnGlnGluLeuGlnGlyVala 210
 QY 657 -----GAGGTGGGCACTTCATGGAGGAGGACGTCCGGATATAC 695
 DB 211 SerTrpGlyTyrGlyCysAlaGluArgAspHis-----ProGlyValTyr 225
 QY 696 ACCAATGTTTACAATATGATCTCGATGATGAGAACACT 734
 DB 226 AlaArgValCysIlePheIleAspTrpLeuGluThrThr 238

RESULT 13

Q9W7Q6 PRELIMINARY; PRT; 238 AA.
 AC Q9W7Q6;
 DT 01-NOV-1999 (Trembl) 12, Created
 DT 01-NOV-1999 (Trembl) 12, Last sequence update
 DT 01-MAR-2003 (Trembl) 23, Last annotation update
 DE Trypsinogen 2 (Fragment).

OS Paratichthys olivaceus (flounder).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Pleuronectiformes;
 OC Pleuronectoidae; Paratichthyidae; Paratichthys.
 OK NCBI_TaxID=8255;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TRISUB=Pancreas;
 RA Suzuki T., Srivastava A.S., Kurokawa T.;
 RT "Japanese flounder mRNA for trypsinogen 2";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB029751; BAA82363.1; -.
 DR HSP; P00763; IDPO.
 DR InterPro; IPRO01254; Ser protease_Try.
 DR Pfam; PF00089; trypsin_1.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS50240; TRYP_SIN_DOM; 1.
 DR PROSITE; PS00134; TRYP_SIN_HIS; 1.
 DR PROSITE; PS00135; TRYP_SIN_SER; 1.
 DR Hydrolyase; Protease; Serine protease.
 FT NON_TFR 1 1
 SQ SEQUENCE 238 AA; 26071 MW; F2B8908085B8D062 CRC64;

Alignment Scores:

Pred. No.:	1.18e-22	Length:	238
Score:	305.50	Matches:	67
Percent Similarity:	52.66%	Conservative:	42
Best Local Similarity:	32.37%	Mismatches:	93
Query Match:	19.58%	Indels:	5
DB:	13	Gaps:	2

US-10-037-270-482 (1-866) x Q9W7Q6 (1-238)

QY 123 CCTATTGGTGAACCTCAAGTCTCACTCAACCCCTGTGGGCGTTCATCAACC 182
 DB 30 ProhibitinValSerLeuAsnSerGlyTyrHisPheCysGlySerLeuValaAnglu 49
 QY 183 AGCTGGGTGGTGGCCCGCCAGCTCACTGCTATTACCAATGGAAGTATGCTGGAAAT 242
 DB 50 AsnTrpValSerAlaAlaHisCysTyrLysSerArgValaGluValaGluMetGlyglu 69
 QY 243 TTCAGAGCAGAGTCAGACGCGTACTGAAACAGCAATTAACCCCATTCAGATGCTCGC 302
 DB 70 HisAsnLeuArgValaGlyGluGlnGlnPheIleSerSerArgValaIlearg 89
 QY 243 TTTCAAGAGCAGAGTCAGACGCGTACTGAAACAGCAATTAACCCCATTCAGATGCTCGC 302
 DB 90 HisProAsnTyrSerSerTyrAsnIleAsnAsnAspIleMetLeuIleLysLeuSerGlu 109
 QY 303 TACTGGAATCACTGATAGCGCCCAAGAGTACCTGATGCTCATCAAGTGGCTAAG 362
 DB 90 HisProAsnTyrSerSerTyrAsnIleAsnAsnAspIleMetLeuIleLysLeuSerGlu 109
 QY 363 CCGACTGTGGGCAAGCTGAGAGGCCCGGTATGCTGTAATTTGGAAGTATTCAGCCGA 422
 DB 110 ProIleThrLeuAsnGlnTyrValaGlnProValaIleLeuProThrSerCysAlaProIle 129
 QY 423 GGCATGCTGTCTACTCTCAAGTGGTGGATGAGCCAGCAAGAAAGTGGCCGACACCT 482
 DB 130 GlyThrMetCysThrValSerGly-----TrpGlyAsnThrMetSerSerThrAspSer 147
 QY 483 GACTTGGGCAAGCTGAGAGGCCCGGTATGCTGTAATTTGGAAGTATTCAGCCGA 542
 DB 148 SerArgGlnGlnCysLeuAspLeuProIleLeuSerGluArgAspCysGluAsnSerTyr 167
 QY 543 CAAGGAAAAGCCACAGAAATTCCTTATGTGTAATTTGGAAGTATTCAGCCGAAT 602
 DB 168 ProGlyMetIleThrAsnAlaMetPheCysAlaGlyTyrLeuGluGlyLysAspSer 187
 QY 603 TTTGGGAGGTGGCGCTGTACTGCTCATCTGCAAGAACAAGTCCAGGGAAATC----- 656
 DB 188 CysGlnGlnLysSerArgLysGlyProValaValCysAsnGlnGluLeuGlnGlyValaSer 207
 QY 657 ---GAGGTGGGCACTTCATGGAGGAGGACGTCCGGATATACCAATAT 713

